2017-06

Analysis of the global maritime transportation system and its resilience

Funk, Daniel

Monterey, California: Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

ANALYSIS OF THE GLOBAL MARITIME TRANSPORTATION SYSTEM AND ITS RESILIENCE

by

Daniel Funk

June 2017

Thesis Co-Advisors:	David L. Alderson
	Ralucca Gera
Second Reader:	Stephen Flynn

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704–0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	06-16-2017	Master's Thesis 01-01-2017 to 06-16-2017

4. TITLE AND SUBTITLE

ANALYSIS OF THE GLOBAL MARITIME TRANSPORTATION SYSTEM AND ITS RESILIENCE

5. FUNDING NUMBERS

6. AUTHOR(S)

Daniel Funk

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

The views expressed in this document are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol Number: N/A.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

The global maritime transportation system carries more than 90% of the foreign trade of the U.S. and many other industrial nations. The loss of a port or the blockade of a canal can cause serious economic consequences, particularly when prearranged deliveries cannot reach their destinations or have to take long detours. Therefore, the efficiency and operability of the global economy highly depends on the resilience of this system.

In this thesis, we analyze the maritime transportation system as a network consisting of container ports, maritime chokepoints and transportation routes between them. We apply the methods and metrics of the network science to find the most central nodes. Furthermore, we formulate a multi-commodity linear optimization model and perform an analysis on different scenarios involving the interdiction of one or more container ports or chokepoints. We contrast the results afforded by the two perspectives. In addition, we evaluate the potential effect of increasing capacity on the arctic sea routes.

14. SUBJECT TERMS

global maritime transportation network, container transport, critical infrastructure, cargo flow, web scraping, network modeling, network analysis, multiplex network, network centrality, multi-commodity linear optimization, network interdiction

15. NUMBER OF PAGES 95

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

ANALYSIS OF THE GLOBAL MARITIME TRANSPORTATION SYSTEM AND ITS RESILIENCE

Daniel Funk
Major, German Army
Dipl.-Inf., Universität der Bundeswehr München, 2009

Submitted in partial fulfillment of the
requirements for the degrees of

MASTER OF SCIENCE IN OPERATIONS RESEARCH
and
MASTER OF SCIENCE IN APPLIED MATHEMATICS
from the
NAVAL POSTGRADUATE SCHOOL
June 2017

Approved by: David L. Alderson
Thesis Co-Advisor

Ralucca Gera
Thesis Co-Advisor

Stephen Flynn
Second Reader, Northeastern University

Patricia A. Jacobs
Chair, Department of Operations Research

Craig W. Rasmussen
Chair, Department of Applied Mathematics

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The global maritime transportation system carries more than 90% of the foreign trade of the U.S. and many other industrial nations. The loss of a port or the blockade of a canal can cause serious economic consequences, particularly when prearranged deliveries cannot reach their destinations or have to take long detours. Therefore, the efficiency and operability of the global economy highly depends on the resilience of this system.

In this thesis, we analyze the maritime transportation system as a network consisting of container ports, maritime chokepoints and transportation routes between them. We apply the methods and metrics of the network science to find the most central nodes. Furthermore, we formulate a multi-commodity linear optimization model and perform an analysis on different scenarios involving the interdiction of one or more container ports or chokepoints. We contrast the results afforded by the two perspectives. In addition, we evaluate the potential effect of increasing capacity on the arctic sea routes.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

1 Introduction **1**
1.1 Background 1
1.2 Literature Review 2
1.3 Our Problem in Context 5

2 Methodology **7**
2.1 Network Framework 7
2.2 Network Formation 16

3 Analysis I – Network Science Analysis **29**
3.1 Network Metrics and Properties 29
3.2 Centralities 36

4 Analysis II – Network Flows / Resilience **41**
4.1 Linear Program Formulation 41
4.2 Preliminary example network 46
4.3 Global Network 51

5 Summary **65**
5.1 Conclusions 65
5.2 Future Work 66

List of References **69**

Initial Distribution List **73**

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

Figure 1.1 Creation of group nodes with example of the Gulf of Mexico. . . 5

Figure 2.1 Visualized tanker traffic density over two weeks in 2016. 8

Figure 2.2 A selection of global maritime chokepoints. 12

Figure 2.3 A HyperText Markup Language (HTML) element tree. 15

Figure 2.4 Example: parsing a web page. 16

Figure 2.5 Example: HTML excerpt from *CIA World Factbook*. 17

Figure 2.6 Distance calculation by SeaRates.com and Ports.com. 18

Figure 3.1 The sea layer visualized in Google Earth. 30

Figure 3.2 Sea routes network graph. 31

Figure 3.3 Adjacency matrix of the sea layer. 33

Figure 3.4 Degree distribution of the sea layer. 33

Figure 3.5 The road layer visualized in Google Earth. 34

Figure 3.6 Road routes network graph. 35

Figure 3.7 Degree distribution of the road layer. 36

Figure 3.8 Plots of centralities' distributions of the sea and the road layer. . . 37

Figure 4.1 Application of node splitting. 42

Figure 4.2 Northern sea route. 43

Figure 4.3 Cargo flow in the base case of the test network. 48

Figure 4.4 Interdiction of the port of Gioia Tauro. 49

Figure 4.5 Interdiction in the Strait of Gibraltar. 50

Figure 4.6 Cargo flow in the base case of the global transportation network. . 51

Figure 4.7 Comparison of the base case and the Suez Canal failure scenario. 53

Figure 4.8 Comparison of the base case and the Panama Canal failure scenario. 54

Figure 4.9 Comparison of the base case and the Strait of Malacca failure scenario. 56

Figure 4.10 Comparison of the base case and the West Coast ports failure scenario. 59

Figure 4.11 Comparison of the base case and the increased arctic traffic scenario. 62

List of Tables

Table 2.1 Selected container ports for our model. 9

Table 2.2 Unicode Transformation Format (UTF)-8 replacement. 20

Table 2.3 Exports and imports of the countries of the model. 22

Table 2.4 Main export and import partners of the countries. 23

Table 4.1 Example of cargo demands. 44

Table 4.2 Results of the base case in the test network. 49

Table 4.3 Results of the interdiction in Gioia Tauro. 50

Table 4.4 Results of the interdiction in the Strait of Gibraltar. 51

Table 4.5 Results of the base case scenario. 52

Table 4.6 Results of the interdiction of the Suez Canal. 53

Table 4.7 Results of the interdiction of the Panama Canal. 55

Table 4.8 Results of the interdiction of the Strait of Malacca. 55

Table 4.9 Nodes with the highest transportation cost increase. 57

Table 4.10 Single-node interdictions resulting in cargo shortfalls. 57

Table 4.11 Ranking of the most “expensive” and the most “central” nodes. . . 58

Table 4.12 Results of the interdiction of the West Coast ports. 60

Table 4.13 Highest cost increase for double interdictions. 61

Table 4.14 Results of the increased capacity in the arctic sea routes. 63

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

AIS	automatic identification system
CIA	Central Intelligence Agency
HTML	HyperText Markup Language
ISO	International Organization for Standardization
nmi	nautical miles
NPS	Naval Postgraduate School
TSI	transportation security incident
TEU	twenty-foot equivalent units
URL	Uniform Resource Locator
UTF	Unicode Transformation Format
UNCTAD	United Nations Conference on Trade and Development
VB	Visual Basic
VBA	Visual Basic for Applications
WTO	World Trade Organization
XML	Extensible Markup Language

THIS PAGE INTENTIONALLY LEFT BLANK

Executive Summary

The global economy is highly dependent on the efficient and reliable transportation of cargo containers. The movement of these containers is facilitated by a global maritime transportation network that consists of ports, waterways, and landside connections.

The global maritime transportation network has been designed to minimize transportation costs and is now highly optimized for efficiency. As long as there are no disruptions to the global system, all cargo sent through the system arrives at its destination without delay. But unforeseen events, like container port disruptions (either deliberate or unintentional) or maritime chokepoint blockades, can force the shipping companies to choose different routes for the cargo, resulting in higher costs or perhaps even making it even impossible for some cargo to be delivered.

This thesis aims to find the weak points of the global network that, if disrupted, may cause the most economic damage due within the system. To reach this goal, we consider two perspectives on the global maritime transportation network—one informed by the connectivity of the global network and another that additionally considers the demands, distances, and capacities that constrain cargo movement.

Our analysis begins with gathering the required data from different sources and combining it to create a network abstraction of the global maritime transportation network. The global maritime transportation network contains hundreds of container seaports worldwide, but many of these are small and do not handle much cargo. In this thesis, we focus on the most important container ports of the world, measured in terms of the amount of cargo throughput. We also include maritime chokepoints—places where sea traffic is naturally constrained—such as straits and canals. In total, we include 94 container seaports from 58 different countries, along with another 26 maritime chokepoints. Although our primary interest is in the movement of cargo by sea, we also consider potential movement of cargo by land, as an alternative in situations where sea transport might be restricted. Thus, we construct a multiplex network model with two "layers" one representing sea transport and another representing land transport. There are a total of 1,518 edges in the sea layer, and 356 edges in the road layer. The edges in our multiplex network are weighted by the real

world travelling distance needed to move the cargo. To measure the total transportation cost in each scenario, the distances are multiplied by the cost per container mile, which is higher for transportation over land than sea. In addition to the distance data, we collect data about the current port throughputs based on the export and import partners of the countries that are represented in the model. Using this data, we derive approximate cargo flows between different pairs of ports.

Within the framework of the network science analysis approach, we study the structure of the sea layer and the road layer and examine them separately. We observe the node distributions, clustering coefficients and the community structure. Using different centrality measurements for each layer, we find that the most central nodes within the network. Within the sea layer, there are the Strait of Gibraltar, the Suez Canal, Bab-el-Mandeb and the Strait of Malacca. These are the chokepoints on the shortest route between Asia and Europe. Within the road layer, the most central nodes are Ambarli, Karachi and Shenzhen. Two of these are also waypoint between Asia and Europe.

We augment this analysis by creating a multi-commodity linear optimization model that minimizes the total cost of the global cargo flow on the basis of our collected data. This model preserves the balance of flow at each node and allows interdictions to single or multiple nodes. We utilize the collected data about cargo flows and translate it into demands for cargo at each node. We also derive from the data the maximum capacities for the ports. We evaluate a base case scenario without any interdictions to the model and measure performance in terms of the total global transportation cost, the fraction of cargo travelling via road transportation, and the number of ports that exhausted their capacities in each scenario. In case of node interdictions, the model re-routes cargo to meet the demands of container ports, where possible. We evaluate the performance of the system for a variety of interdiction scenarios involving the closure of one or two nodes, and we rank the nodes based on the increased costs to the system. This perspective of the analysis shows that closure of the port of Busan (South Korea) causes the biggest increase of transportation cost due to the additional road transportation of cargo to the closest ports in the area. Another expensive scenario would be the blockade of the Strait of Gibraltar, which would cause long detours for the cargo vessels. From all possible scenarios of double port interdictions, the simultaneous closure of the ports in Los Angeles and Long Beach would cause by far the most economic damage. Other expensive scenarios mostly contain closures of ports

in Asia, but scenarios with closures of ports on different continents can also result in big increase of transportation cost.

As a final step, we compare both rankings of the most central nodes (based on betweenness centrality) and the costly nodes (whose failure produce the greatest transportation cost increase). We conclude that the techniques of network science provide reasonable candidates for the costly nodes of our network, but not a complete list of the top costly ones. To assess the impact of port closures or blockades, it is necessary to include not only connectivity, but also capacities, costs, and demands.

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

First of all, I would like to thank the German Army for giving me the incredible chance to study at the Naval Postgraduate School. The knowledge I have gained here will enrich my professional life and help me to contribute toward enhancing the German Armed Forces.

I also would like to thank my thesis advisors, Professor David L. Alderson and Professor Ralucca Gera, for the great support during the entire time. Working on this thesis with their support and receiving professional insights from two different fields of specialization improved the outcome of this work and filled my time at NPS with many beneficial experiences.

Most important, I would like to thank my beloved girlfriend, who showed incredible patience and never lost her faith in me. It would have not been possible to get through the endeavor of this two-year study without her constant support.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

1.1 Background

Increased globalization over the last few decades has caused a huge increase in international commodity trade. According to the World Trade Organization (WTO) the growth in world trade increased from 1950 to 2005 by a factor of 27 [1]. Outsourcing of production and lower labor costs in certain parts of the world have decreased manufacturing costs and therefore, the profit margin for many products. Simultaneously, there has been a growing need for transportation of more and more goods across the world at cheaper costs.

This gap was filled in the late 1950s, when the International Standards Organization (ISO) container was introduced and revolutionized global shipping. The standardization of maritime transport, through the use of containers and increased mechanization of the container handling in seaports, reduced the needed manpower by about 90% and the container handling cost by about 80% [2]. The global maritime transportation network took advantage of that development and grew very quickly. Within this global network, some seaports had logistics advantages—for example, convenient geographical location or better developed landside transport connections—and therefore, they were favored by the transportation companies and became regional hub ports. In a manner consistent with the "rich get richer" principle, these few ports grew faster than the others, became global "megaports" and are today essential for the worldwide logistics.

Today, nearly a half of the worldwide container handling is transshipped through the top twenty global megaports [3]. Therefore, today's megaports are not only important for the surrounding regions, but also for the entire global maritime transportation network due to potential vulnerabilities that they might create. It is widely believed that the loss of one or more megaports—as could happen from a terrorist attack, piracy actions, infrastructure failure, or mere capacity limitations—could have severe global consequences. One famous example is the West Coast port labor slowdown/lockout in 2002, as negotiations of labor agreements caused a stop of port operations for ten days. As *Los Angeles Times* reports:

"It took the West Coast ports 100 days to return to normal operations" [4]. A longer port shutdown would have caused a shutdown of production lines in the United States and emptied the shelves in the malls [5].

1.2 Literature Review

Many studies model and analyze different transportation networks. They usually differ in many aspects, like the type and size of the modeled network, its granularity, and even the measurement of performance.

1.2.1 The Network Science Perspective

Within the growing field of network science, transportation networks are a common topic of study for analyzing network structure, performance, resilience and other measures of interest. Many kinds of networks can be examined: public transportation of passengers, transport of goods or even computer networks like the Internet. They can consist of one layer or of multiple layers as multiplex networks.

As an example, Guimerà et al. [6] study the structure of the worldwide air transportation network. They analyze whether this network has a scale-free and/or small-world network structure, and they observe that the most "central" cities are not necessarily the most-connected. Furthermore, they discover that the network can be subdivided in certain communities and they identify these.

Another example by Blasuis et al. [7] considers the global maritime transportation network in a manner similar to our problem. They aggregate transportation flows of goods by type (i.e., bulk dry carriers, container ships and oil tankers) as well as by their specific physical characteristics. Each of these three types of ships is represented as its own layer, so the network ends up with a multi-layered structure. The authors analyze the ship movements, gained from their automatic identification system (AIS) transmitters, to understand patterns of global trade and bioinvasion.

Wang and Wang present the spatial pattern of the global maritime transportation network [8]. They reveal that the worldwide system is subdivided in 44 regional hub-and-spoke subsystems with a hierarchical structure. Their model is based on raw monthly schedules of 24

shipping carriers. “East Asia, Southeast Asia, Northeast Europe, and East coast of the USA are the concentration regions of worldwide shipping lines” [8]. The paper demonstrates a great diversity in linkage coverage (probability that one port is connected with any other port) among different regions.

1.2.2 The Operations Research Perspective

Martagan, Eksioglu, Eksioglu and Greenwood [9] “develop a simulation model that can be used to make effective re-routing decisions so that the time for freight to reach its final destination is not significantly increased in a crisis.” This simulation model of the maritime transportation network is developed for macro-level analyzes. As the main performance measure, they take into account the lead time for freight from its origin to its destination. In the study, they collect data from simulation runs and evaluate the performance of the model, first without disruptions and then with various predefined disruption scenarios. Finally, they apply and compare some re-routing strategies for ships in the model. The model is created by a probabilistic approach and it works with different probability distributions for the time how long ships stay in a port, number of containers per ship, distribution of container destinations, etc. However, their study is limited to the U.S. and includes only seven ports.

Madhusudan and Ganapathy [10] study the disaster resilience of transportation infrastructure and ports. They first give an overview on the general topic of disasters and later, in particular, on disaster resilience of transportation infrastructure and transportation networks. Finally, they specifically examine port resilience, stating that seaports are vital to most nations' economies. The paper introduces different useful measurements of resiliency.

Miller-Hooks et al. [11] examine “the problem of measuring a network's maximum resilience level and simultaneously determining the optimal set of preparedness and recovery actions necessary to achieve this level.” The network in the model is an abstraction of the U.S. rail-based intermodal container network. The problem of measuring resilience given preparedness options is formulated as a two-stage stochastic program, including the preparedness actions before the disaster and the recovery actions after it. The goal of the paper is to find the best portion of the budget to spend on preparedness and the portion to save for the recovery actions after the disaster. They solve this for a number of predefined

possible scenarios.

1.2.3 Previous Work at NPS

Previous work at the Naval Postgraduate School (NPS) has considered the impact of port closures on the flow of commodities.

Pidgeon [12] models the seven major ports of the U.S. West Coast to evaluate disruptions and costs inflicted on the shipping industry by a transportation security incident (TSI). He examines the potential bottlenecks within the port infrastructure, that can be vulnerable to a TSI and reveals the threats for the domestic cargo-handling capacity. The thesis evaluates different predefined scenarios of infrastructure interruptions and their impact on the economy and provides recommendations future investments to alleviate port congestion.

Bencomo [13] undertakes similar research to Pidgeon's thesis and aims to find critical infrastructure components, but he includes the entire U.S. for his model. The study implements the data about international container flows between 46 countries and the U.S. ports. The cargo transportation contains different commodities, where each commodity is an origin-destination flow. The thesis utilizes an Attacker-Defender model, introduced by Brown et al. [14], [15], to first initiate interdictions to the network and then lead the cargo to its destinations through the surviving part of the network.

Babick [16] studies a new approach to utilize a Design-Attack-Defend model [14], [15] to determine an optimal defense plan for critical infrastructure. Based on the U.S. West Coast rail system, the model first identifies a worst case scenario attack and then determines where to defend the system or to build additional infrastructure in order to minimize the impact of the attack. The model works with budget constraints that limit the possible enhancements of the optimal defense plan.

Onuska [17] models the coal transportation network in the Port of Pittsburgh to study different "what if" scenarios. The goal of the thesis is to determine the critical infrastructure and to assess the resiliency of the system. With rivers, rails and roads, the network contains three different modes of transportation. As some previous works, this thesis applies Defender-Attacker-Defender techniques [14], [15] to provide an optimal defense plan for most likely scenarios.

Garcia Olalla [18] models the global maritime transportation network with a large selection of seaports from all continents. To model and simplify the connections between the seaports, he creates artificial transition points (group nodes), which bundle all connections to and from certain regions with multiple seaports. An example of the Gulf of Mexico is displayed in Figure 1.1. Most of these group nodes are connected to others, which results in a type of "maritime highway." To stress the model and study its resilience, he considers disruptions of the edges, which represent the sea connections between seaports. First, he analyzes scenarios with one disruption and then with disruptions on multiple connections. The disruptions are expressed by either time or cost penalties for the ships that have to take the penalized route, if their route is affected. Our thesis builds on his work, by extending and refining some of his approaches.

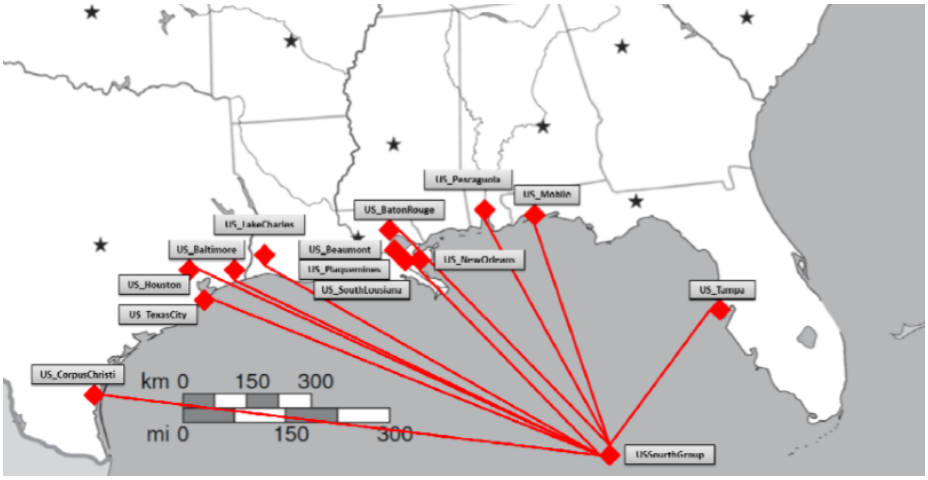


Figure 1.1. Garcia Olalla's creation of group nodes with example of the Gulf of Mexico. Source: [18].

Some of the previous work at NPS is not available to this researcher due to its restricted nature. This includes an assessment of the resilience of the Port of Honolulu [19] and the Port of Anchorage [20].

1.3 Our Problem in Context

In this thesis, we analyze the global maritime transportation network and its resilience from two perspectives. First, we apply the techniques of network science to find the most "central" container ports or transition points within the network. We also examine its community

structure and try to explore potential anomalies. Then, we adopt an operations research perspective and present a network flow model that also considers distances, capacities, and costs in the global delivery of container goods.

Within the context of the previous problems and modeling approaches, our work makes the following contributions:

- We model the transportation network as a global network, including all continents, instead of a regional excerpt.
- We construct our network as a multiplex network with one layer for the shipping network and one for road transportation along the landside connections.
- We only concentrate on transportation of containers. Potentially valuable insights from oil and bulk transportation are not covered by this thesis.
- We create our model, using real sea and land connection distances between ports instead of using approximations.
- We do not single ship movements, per se, but we observe the movements of goods between ports on the shortest/cheapest available route.
- We establish re-routing strategies that apply, if a part of a route becomes impassable for container ships.

We analyze the global impact of the loss of one or more container ports. Using the base case of no disruptions, we measure the amount of goods that have to be re-routed in case of each disruption. We analyze the effect of single disruptions as well as multiple disruptions on global transportation.

CHAPTER 2:
Methodology

This chapter introduces an overview of the initial work before the actual analysis. This includes gathering the required data from different sources and combining it to create a network abstraction of the global maritime transportation network.

2.1 Network Framework

We create a network model of the global maritime transportation network to be studied from different perspectives. Our goal is to quantify its resilience in the presence of local or regional disruptions. We seek a model that reflects the most important characteristics of the maritime network.

The global maritime transportation network, particularly the portion that carries container traffic, consists of container ships of different sizes, seaports around the world, and finally, the chokepoints of the maritime network. In this thesis, a seaport implies the ability to move container cargo between container ships and land.

Rodrigue [21] describes maritime chokepoints as follows:

> Chokepoints are a common concept in transport geography, as they refer to locations that limit the capacity of circulation and cannot be easily bypassed, if at all. This implies that any alternative to a chokepoint involves a level of detour or use of an alternative that translates into significant financial costs and delays.

A chokepoint can be a natural place where ship traffic is condensed by the geographic obstacles, like the Strait of Gibraltar, Strait of Malacca or the Strait of Magellan. A chokepoint can also be a man-made passage, which allows ships to cross the landmass through artificial canals in order to shorten their routes to the destination port. Examples of such chokepoints include the Panama Canal or the Suez Canal. In addition, we will count an area as a maritime chokepoint even if it is not a narrow but contains concentrated

ship traffic and is a waypoint on many long distance sea routes. One example of such a chokepoint is the Cape of Good Hope, which is still a very frequented area nowadays, as seen in Figure 2.1.

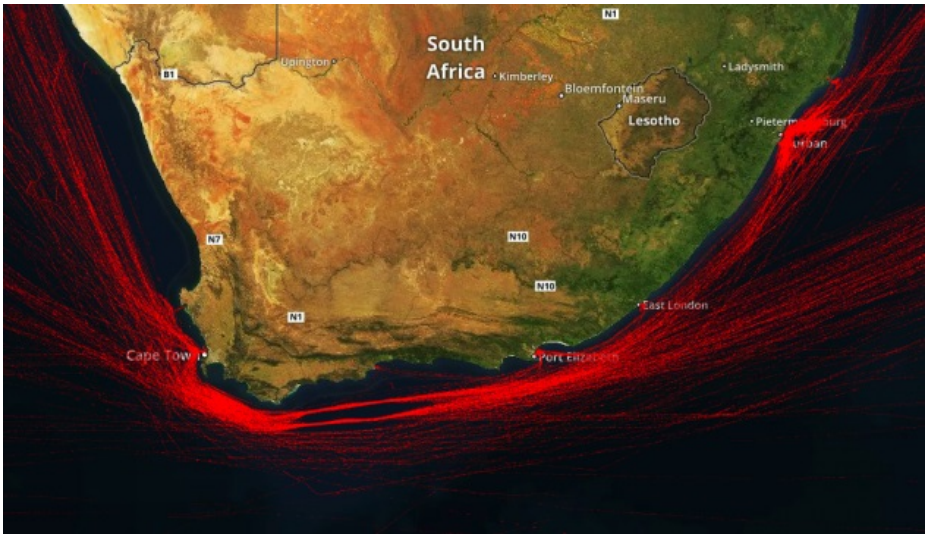


Figure 2.1. Visualized tanker traffic density at the Cape of Good Hope over two weeks in 2016. Source: [22].

2.1.1 Selection of the Ports

The global maritime transportation network contains hundreds of container seaports worldwide, but many of these are small and do not handle much cargo. In this thesis, we focus on the most important container ports of the world, measured in terms of the amount of cargo throughput. This provides a reasonable representation of global maritime transportation, while still including ports from nearly every continent.

The Central Intelligence Agency (CIA) produces an annual *World Factbook*, which contains various statistical data about the countries of the world. Among other versions, it is published as a permanently updated online version [23]. One section there is called "Transportation" and provides data about major container ports of many countries, including the throughput measured in terms of twenty-foot equivalent units (TEU). The currently available throughput data is mostly from 2011, with few exceptions of 2010 and 2012. In total, there are 94 container seaports from 58 different countries. We use these 94 ports for our global model. A summary of the ports is displayed in Table 2.1.

Table 2.1. Selected container ports for our model. Source: [23]

Country	Port	Throughput (in TEU)
Argentina	Buenos Aires	1,851,701
Australia	Brisbane	1,004,983
Australia	Melbourne	2,467,967
Australia	Sydney	2,028,074
Bahamas	Freeport	1,116,272
Bangladesh	Chittagong	1,392,104
Belgium	Antwerp	8,664,243
Belgium	Zeebrugge	2,207,257
Brazil	Itajai	983,985
Brazil	Santos	2,985,922
Canada	Metro Vancouver	2,507,032
Canada	Montreal	1,362,975
China	Dalian	6,400,300
China	Ningbo	14,719,200
China	Qingdao	13,020,100
China	Port of Shanghai	31,739,000
China	Tianjin	11,587,600
China	Guangzhou	14,260,400
China	Shenzhen	22,570,800
Colombia	Cartagena	1,853,342
Ecuador	Guayaquil	1,405,762
Egypt	Alexandria	1,108,826
Egypt	Port Said	3,755,796
France	Le Havre	2,215,262
Germany	Bremerhaven	5,915,487
Germany	Hamburg	9,014,165
India	Chennai	1,558,343
India	Jawaharlal Nehru Port	4,307,622
Indonesia	Tanjung Priok	5,617,562

Continued on next page

Country	Port	Throughput (in TEU)
Iran	Bandar Abbas	2,752,460
Ireland	Dublin	1,931,001
Israel	Ashdod	1,176,000
Israel	Haifa	1,238,000
Italy	Genoa	1,847,648
Italy	Gioia Tauro	2,264,798
Italy	La Spezia	1,307,274
Jamaica	Kingston	1,724,928
Japan	Kobe	2,725,304
Japan	Nagoya	2,471,821
Japan	Osaka	2,172,797
Japan	Tokyo	4,416,119
Japan	Yokohama	2,992,517
Korea, South	Busan	16,163,842
Korea, South	Kwangyang	2,061,958
Korea, South	Incheon	1,924,644
Lebanon	Beirut	1,034,249
Malaysia	Penang	1,202,180
Malaysia	Port Klang	9,435,403
Malaysia	Tanjung Pelepas	7,302,461
Malta	Marsaxlokk	2,360,000
Mexico	Manzanillo	1,992,176
Mexico	Lazaro Cardenas	1,242,777
Morocco	Tangier	2,093,408
Netherlands	Rotterdam	11,876,920
Oman	Salalah	3,200,000
Pakistan	Karachi	1,545,434
Panama	Balboa	3,232,265
Panama	Colon	2,390,976
Panama	Manzanillo International Terminal	2,391,066
Peru	Callao	1,616,365

Continued on next page

Country	Port	Throughput (in TEU)
Philippines	Manila	3,342,200
Puerto Rico	San Juan	1,484,595
Russia	Saint Petersburg	2,365,174
Saudi Arabia	Jeddah	4,010,448
Saudi Arabia	King Abdul Aziz Port	1,492,315
Singapore	Singapore	31,649,400
South Africa	Durban	2,712,975
Spain	Las Palmas	1,287,389
Spain	Algeciras	3,608,301
Spain	Barcelona	2,033,747
Spain	Valencia	4,327,371
Sri Lanka	Colombo	3,651,963
Taiwan	Keelung	1,749,388
Taiwan	Kaohsiung	9,363,289
Taiwan	Taichung	1,383,578
Thailand	Bangkok	1,305,229
Thailand	Laem Chabang	5,731,063
Turkey	Mersin	1,126,866
Turkey	Ambarli	2,121,549
United Arab Emirates	Dubai	12,617,595
United Arab Emirates	Khor Fakkan	3,234,101
United Kingdom	Southampton	1,324,581
United Kingdom	Felixstowe	3,248,592
United Kingdom	London	1,932,000
United States	Long Beach	6,061,091
United States	Los Angeles	7,940,511
United States	Oakland	2,342,504
United States	Seattle	2,033,535
United States	Houston	1,866,450
United States	New York	5,503,485
United States	Savannah	2,944,678

Continued on next page

Country	Port	Throughput (in TEU)
United States	Hampton Roads	1,918,029
Vietnam	Hai Phong	1,018,794
Vietnam	Saigon Port	3,071,777

2.1.2 Selection of the Chokepoints

Chokepoints represent areas with restricted throughput and/or high concentration of ships, which makes them very important and potentially vulnerable points of the whole network. Figure 2.2 shows the main maritime transportation routes and identifies several chokepoints along the core routes.

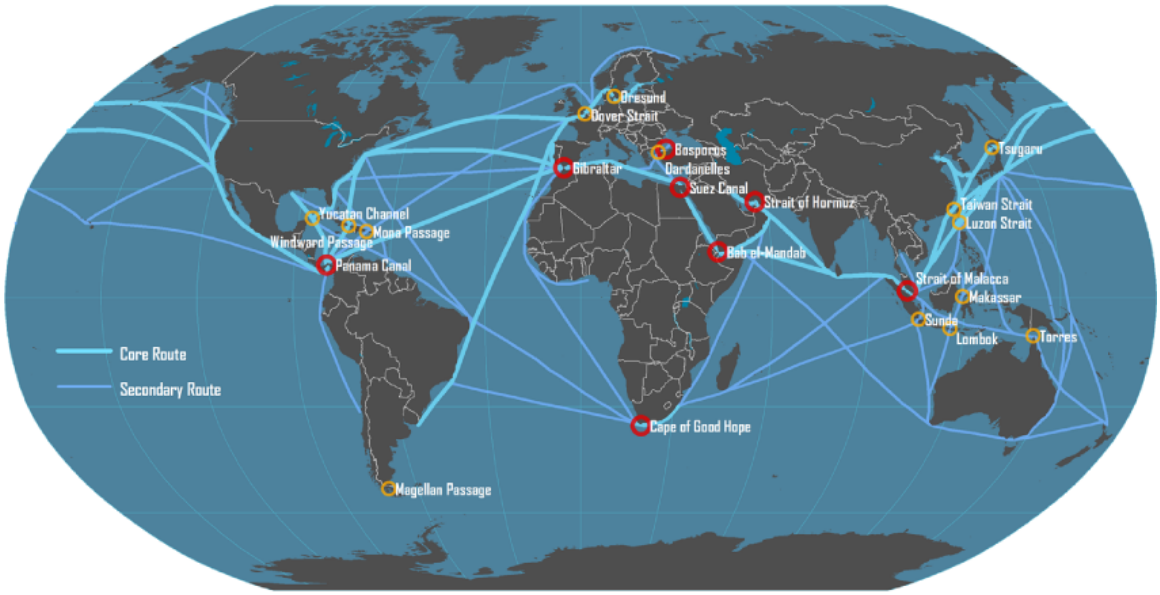


Figure 2.2. A selection of global maritime chokepoints. Source: [24].

There are many places throughout the world that could potentially be considered maritime chokepoints, however, only a small subset of representative chokepoints is required for our network model.

Komiss and Huntzinger [25] choose maritime chokepoints based on established oil tanker transportation routes. These are mostly located on the main routes from the Middle East and.include the following: the Strait of Hormuz, the Strait of Malacca, the Suez Canal,

Bab el-Magdeb, the Turkish Straits (Dardanelles), and the Panama Canal. Noer et al. [26] considers chokepoints of the trade routes and strategic straits in the Australasian Mediterranean Sea, which includes the Lombok Strait and the Sunda Strait. Other considerations for chokepoints take into account less frequented but still important chokepoints like the Magellan Passage or the Dover Strait.

Based on these considerations, along with our previous selection of the container ports, we choose 26 maritime chokepoints, which cover the routes between our selected ports. These are the following:

- Bering Strait,
- Davis Strait,
- Barents Sea,
- Strait of Hormuz,
- Strait of Malacca,
- Bab-el-Mandeb,
- Panama Canal,
- Suez Canal,
- Strait of Gibraltar,
- Cape of Good Hope,
- Sunda Strait,
- Lombok Strait,
- Torres Strait,
- Makassar Strait,
- Taiwan Strait,
- Luzon Strait,
- Magellan Passage,
- Dardanelles,
- Dover Strait,
- Øresund,
- Great Australian Bight,
- the northern tip of Great Britain,
- the northern tip of Trinidad and Tobago,
- Yucatan Channel,

- Windward Passage, and
- Mona Passage.

2.1.3 Data Collection

To the best of our knowledge, there does not exist a standard representation for the global maritime transportation network nor is there is a single repository from which one can obtain complete data to support its creation. Therefore, an essential part of our work is the collection of underlying data from selected websites.

Because we need to collect a large amount of information housed at a variety of different websites, we employ "web scraping" (also known as "screen scraping" or "web harvesting") to automatically gather data from the Internet. Mitchell [27] describes web scraping as follows:

> In theory, web scraping is the practice of gathering data through any means other than a program interacting with an [Application Program Interface] (or, obviously, through a human using a web browser). This is most commonly accomplished by writing an automated program that queries a web server, requests data (usually in the form of the [HyperText Markup Language] and other files that comprise web pages), and then parses that data to extract needed information.

There are several programming languages which offers tools for web scraping. The basic idea underlying such tools is to obtain the underlying code in HyperText Markup Language (HTML) for the website of interest using its Uniform Resource Locator (URL). This code can be stored as text and then searched for the relevant data. Provided some knowledge of HTML programming and the website of interest, one can gather the required information from the resulting code using simple text search.

The underlying code for a website is based on elements that form the building blocks of HTML pages, containing tables, images, text modules and other components of websites. These HTML elements are delimited by "tags" denoted by angle brackets, like `<h1>...</h1>` or `<b>...</b>`, that are organized in a tree structure. Figure 2.3 represents an example of such an HTML element tree. A number of software libraries are

available help to automate the parsing of such tags, but they require that the user knows between which tags the desired data is stored.

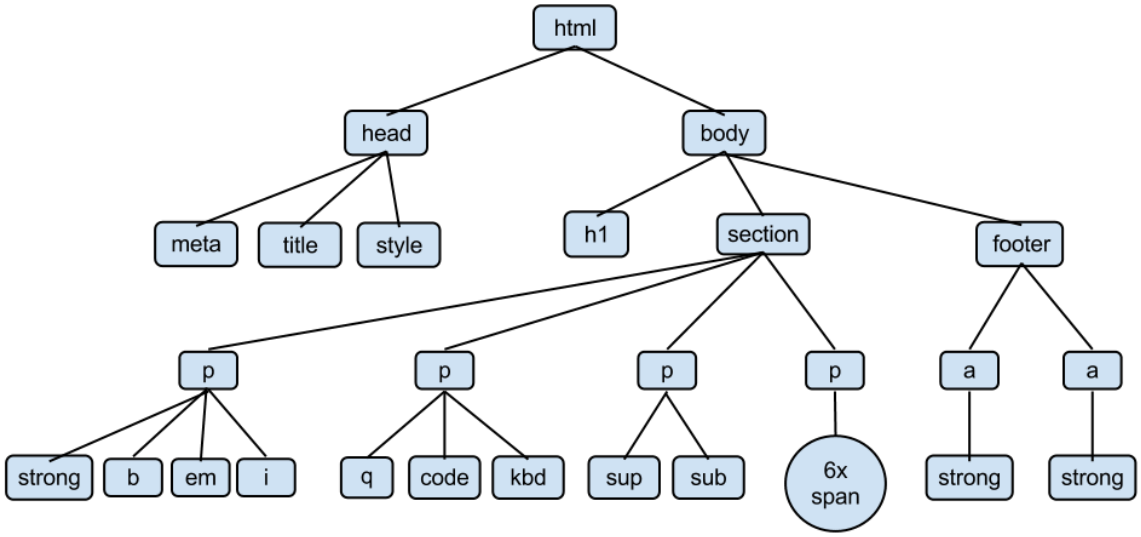


Figure 2.3. A HTML element tree. Source: [28].

For the data collection in this thesis, we choose to work with Visual Basic for Applications (VBA), which is closely related to Visual Basic (VB) but is only able to run within the Microsoft Office applications. One of the advantages to collect the data using VBA within Microsoft Excel is that the collected data can be easily stored, examined and processed directly in a worksheet.

VBA provides multiple means of web scraping. The most simple way is using Internet Explorer controlled by VBA. Figure 2.4 displays code that will open a website in Internet Explorer and store its source code in a variable named `html`. The `Do While` loop makes sure that the algorithm will wait until the website is completely loaded in the web browser before starting accessing it. Further parts of the code, like closing Internet Explorer, have been removed for a better clarity of the example.

```
Sub GetWebsite()
    Dim ie As InternetExplorer
    Dim html As HTMLDocument

    Set ie = New InternetExplorer
    ie.Visible = True
    ie.navigate "https://www.unibw.de/"

    Do While ie.readyState <> READYSTATE_COMPLETE
        DoEvents
    Loop
    Set html = ie.document
End Sub
```

Figure 2.4. VBA code to open and parse a website in Internet Explorer.

Finding the right position of the desired data in the HTML code can be challenging when it is done by an algorithm. The algorithm needs to know by which tag the data is surrounded. Some HTML tags contain additional attributes, like `id` or `class`. They are included in the source code. This can be very helpful when the right tag that needs to be found is just one of many with the same tag name. Figure 2.5 displays HTML code corresponding to an extract of the *CIA World Factbook* website [23], listing ports and terminals of the countries of the world. The explicit URL is https://www.cia.gov/library/publications/the-world-factbook/fields/2120.html. The website organizes the entries of the countries in a long table. The code below only represents the entry for the German ports. The `tr` tag stands for the table row and `td` tags are the table cells. After studying the entire document, one realizes that the `tr` tag IDs mean the country of the entry. Here, `gm` stands for Germany. Knowing this, it is easy to get to the desired country. Furthermore, one needs to recognize that `fieldData` is the class name of every `td` where the container port data is stored, which is exactly what we need. Therefore, we can establish an iterative algorithm which searches the World Factbook and finds every sea container port, and its annual throughput (in TEU), in every country.

2.2 Network Formation

Although our primary interest is in the movement of cargo by sea, we also consider potential movement of cargo by land, as an alternative in situations where sea transport might be

```
<tr id=gm>
    <td class=country>
        <a href=../geos/gm.html>Germany</a>
    </td>
    <td class=fieldData>
        <strong>container port(s): </strong>
        Bremen/Bremerhaven (5,915,487), Hamburg (9,014,165) (2011)
        <br />
        <strong>LNG terminal(s) (import): </strong>
        Hamburg
        <br />
    </td>
</tr>
```

Figure 2.5. Parsing this type of data allows us to automate the collection of port throughput data.

restricted. Thus, our network model has two "layers" one representing sea transport and another representing land transport. In general, transportation by rail is preferred over transportation by road because it is lower cost, however rail coverage for each continent is not universal. For this reason, we choose to represent road transport as the additional layer.

Given the selection of network nodes as described above, we need to identify the corresponding network edges. The following subsections describe the process for selecting the edges for both layers and of the collection of corresponding data.

2.2.1 Calculation of Distances

The edges of our model represent the routes between container ports (or chokepoints). To make the model realistic, the edges need to have a weight attribute that represents how "expensive" it is to use the edge on a route. Usually, this is expressed by the travelling time or the actual transportation cost, both of which are usually based on the distance between nodes.

Therefore, our model requires the distance between each node. To provide a more accurate analysis we collect real data of the sea and road distances, instead of estimating it. We use web scraping, in a manner similar to that described above, to facilitate the data collection. We record all of our calculated distances in nautical miles (nmi) so the network will be

consistent for further analysis.

Sea Distances

There are a large number of websites that provide sea route distance calculations. It is difficult to see how accurate the calculations are. Some websites provide only the resulting distance, while others show the calculated route on a map. This visualization often helps to see if the chosen route is reasonable or is a result of a poor algorithm. As an example, Figure 2.6 demonstrates two calculations from different websites of a sea route from Miami to Rotterdam. While the left figure shows the direct route, the right one works with horizontal movements making many detours. The actual results of the calculations show the real difference – 4,067nmi in the left figure and 5,009nmi in the right figure. A sanity check, done by hand, shows, that the calculations of the left website provide indeed much better results than the right website. Initially, we use SEA-DISTANCES.org, a website with more than 4,000 seaports, according to its own statement. It provides valid calculations and multiple routes between two ports. Unfortunately, not all of our 94 chosen ports for the model are included on this website. Therefore, we decide to choose SeaRates.com for our calculations and data collection.

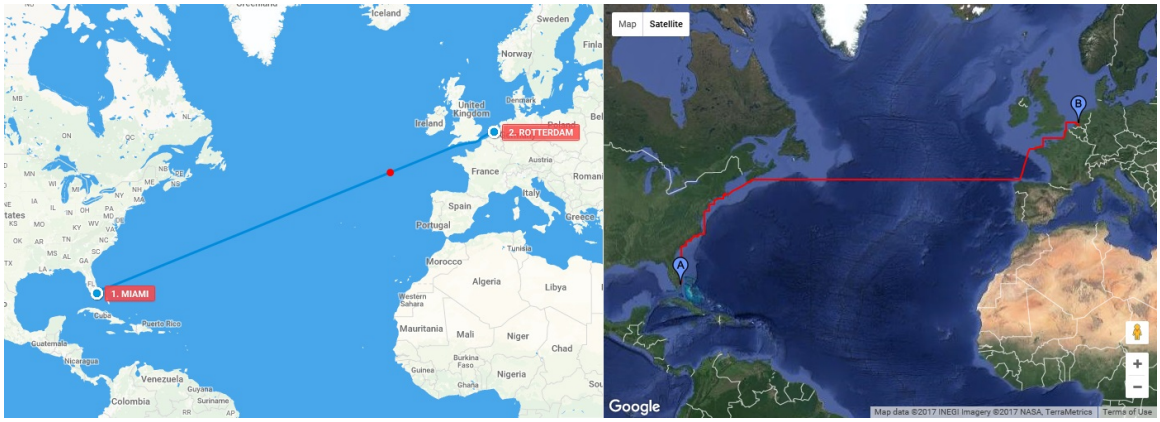


Figure 2.6. Distance calculation by SeaRates.com (left) and Ports.com (right).

We use an adjacency matrix to store the distance data. It has zeros on the main diagonal and is symmetric, since we assume that the direction of a route between two nodes doesn't

influence the actual travelling distance. Because we assume that to travel to a different body of water ships have to go through a chokepoint, most of the entries in the adjacency matrix are `NA`. There are only connections within the same body of water and to the next chokepoints. Only those distances are needed for the model.

To conduct web scraping for our purposes, we use https://www.searates.com/reference/portdistance/, which is a subpage of SeaRates.com. The following URL provides a route distance from Lisbon to New York:

https://www.searates.com/reference/portdistance/?D=**11862**&G=**16959**&shipment=1&container=20st&weight=1&product=0&request=&weightcargo=1&

The `D` and `G` attributes in the URL are the origin and the destination of the route. Every port has its own ID in SeaRates.com, so we first extract the port IDs for our ports from the website. Once it is done, we simply open the correct URL to get the route calculation. From the resulting website we extract the route distance with the earlier described technique. The only challenge for the sea route layer of our network is the calculation of distances to or from a chokepoint. To do this with an algorithm using SeaRates.com, we find container ports, which are close to the chokepoints and available on the website. Then we use these ports for the calculations, which we assume to be a good approximation of the real route distance. Whenever there are no close ports available, which is a very rare case, we do the route distance calculations by hand with another website http://www.sea-seek.com/tools/tools.php.

Road Distances

For the road route distances, we create a new adjacency matrix, which is very similar to the sea route matrix. Since the chokepoints are only relevant for the sea layer, we do not include them in the road matrix. From SeaRates.com, we collect the coordinates of the chosen container ports. These are not as important for the sea route distance calculations, but they become important for the road route distance calculations and visualization of the model. The coordinates of the ports enable us to collect the road distances from Bing Maps and Google Maps websites. As an example for Bing Maps we use the following URL to get the distance between two ports:

http://dev.virtualearth.net/REST/V1/Routes/Driving?o=xml&wp.0=**51.21666667**,**4.4**&
wp.1=**59.93333333**,**30.3**&distanceUnit=km&key=
Anm_HN0shZZaRTHzt1bDAcyzMpOEGt7KUao63aycqlg3l2lY9LZSCCIJCTV1Muws

The bold numbers are the latitudes and longitudes of both ports in decimal degrees. The URL returns an Extensible Markup Language (XML) document that has a similar structure as HTML, which was introduced earlier. The distance can be easily extracted from this document. For some few road routes, Bing Maps either doesn't provide a route calculation or provides a route, which is obviously by far not optimal. For both cases, we fill up the road route adjacency matrix with Google Maps queries, for example using the following URL:

https://www.google.com/maps/dir/
35%C2%B050%E2%80%B230.00%E2%80%B3N+
14%C2%B032%E2%80%B241.00%E2%80%B3E/
53%C2%B032%E2%80%B260.00%E2%80%B3N+
9%C2%B055%E2%80%B260.00%E2%80%B3E

Again, the bold parts represent both coordinates, but here measured in degrees, minutes and seconds. Here an example: 48°08′14″ N, 11°34′31″ E. The for coordinates typical punctuation marks need to be replaced here by Unicode Transformation Format (UTF)-8 bytes as follows in Table 2.2, so they can be interpreted by the Google Maps website:

Table 2.2. UTF-8 replacement.

degree sign	°	C2 B0
prime	′	E2 80 B2
double prime	″	E2 80 B3

For our model, not every road connection between two ports is required. Since road transportation is expensive, it would only be used to bypass relative short distances, compared to those within the sea layer. The purpose of the road layer will be either to deliver goods to the final destination or to the next functional seaport to switch back to ship transportation. Therefore, we set a threshold for the road connections at maximum 2,000nmi. We set longer distances to `NA` in the adjacency matrix.

2.2.2 Flows

The *CIA World Factbook* provides data about exports and imports of the countries of the world. More precisely, the data contains the main trading partner countries with the corresponding percentage and the total trading volume for each country. These specifications are given for exports and imports separately. Therefore, we can reconstruct approximately the outgoing and the incoming flows of goods from and to each country in our model.

To make an example, we choose Germany with an export volume of $1.283 trillion (2016 est.). The main exports partners (2015) are:

- United States 9.6%
- France 8.6%
- United Kingdom 7.5%
- Netherlands 6.6%
- China 6%
- Italy 4.9%
- Austria 4.8%
- Poland 4.4%
- Switzerland 4.2%

The total export volume covered by the main partners is 56.6%, which corresponds to an export volume of $726 trillion. In our model two container ports of Germany are represented. These are Bremerhaven (5,915,487 TEU) and Hamburg (9,014,165 TEU). According to their throughput, we assign a proportional percentage of all numbers for Germany with the distribution: Bremerhaven 40% and Hamburg 60%. For all seaports of one country we assume the same distribution of exports, as for the country itself. Therefore, for example from Bremerhaven, we assume an export flow of 9.6% of the 5,915,487 TEU to the United States. On the other hand, the same flow from Bremerhaven will be distributed as import to the United States ports in the same manner. For the import data we apply the same procedure.

Table 2.3 shows the export and import data from the *CIA World Factbook*, that we use in model.

Table 2.3. Exports and imports of the countries of the model. Source: [23]

Country	Exports	Imports
Argentina	$58.4 billion	$57.23 billion
Australia	$184.3 billion	$203.1 billion
Bahamas	$880 million	$2.495 billion
Bangladesh	$33.32 billion	$39.17 billion
Belgium	$250.8 billion	$251.7 billion
Brazil	$189.7 billion	$143.9 billion
Canada	$402.4 billion	$419 billion
China	$2.011 trillion	$1.437 trillion
Colombia	$33.64 billion	$47.15 billion
Ecuador	$16.77 billion	$17.74 billion
Egypt	$14.73 billion	$50.07 billion
France	$505.4 billion	$525.4 billion
Germany	$1.283 trillion	$987.6 billion
India	$271.6 billion	$402.4 billion
Indonesia	$136.7 billion	$121.5 billion
Iran	$87.52 billion	$62.12 billion
Ireland	$160.1 billion	$88.01 billion
Israel	$51.61 billion	$57.9 billion
Italy	$436.3 billion	$372.2 billion
Jamaica	$1.278 billion	$3.772 billion
Japan	$641.4 billion	$629.8 billion
Korea, South	$509 billion	$405.1 billion
Lebanon	$3.108 billion	$17.98 billion
Malaysia	$167.3 billion	$139.5 billion
Malta	$2.915 billion	$4.479 billion
Mexico	$359.3 billion	$372.8 billion
Morocco	$18.72 billion	$33.15 billion
Netherlands	$460.1 billion	$376.3 billion
Oman	$30.39 billion	$25.78 billion

Continued on next page

Country	Exports	Imports
Pakistan	$20.96 billion	$38.25 billion
Panama	$15.19 billion	$22.08 billion
Peru	$38.09 billion	$38.35 billion
Philippines	$38.2 billion	$60.95 billion
Puerto Rico	$70.41 billion	$47.61 billion
Russia	$259.3 billion	$165.1 billion
Saudi Arabia	$205.3 billion	$157.7 billion
Singapore	$353.3 billion	$271.3 billion
South Africa	$83.16 billion	$85.03 billion
Spain	$266.3 billion	$287.9 billion
Sri Lanka	$10.12 billion	$18.64 billion
Taiwan	$314.8 billion	$248.7 billion
Thailand	$190 billion	$171.3 billion
Turkey	$150.1 billion	$197.8 billion
United Arab Emirates	$316 billion	$246.9 billion
United Kingdom	$412.1 billion	$581.6 billion
United States	$1.471 trillion	$2.205 trillion
Vietnam	$169.2 billion	$161 billion

Table 2.4 shows the export and import partners of each country. The countries without ports in our model have been filtered out. Since there is no data provided for Puerto Rico, we assume the same partners there as for the United States.

Table 2.4. Main export and import partners of the countries. Source: [23]

Country	Export Partners	Import Partners
Argentina	Brazil 17%, China 8.6%, US 5.9%	Brazil 22.4%, US 16.3%, China 15.5%, Germany 5.1%
Australia	China 32.2%, Japan 15.9%, South Korea 7.1%, US 5.4%, India 4.2%	China 23%, US 11.2%, Japan 7.4%, South Korea 5.5%, Thailand 5.1%, Germany 4.6%

Continued on next page

Country	Export Partners	Import Partners
Bahamas	US 15.9%	US 22.3%, China 14.8%, Japan 9.5%, South Korea 7.3%, Colombia 6.8%, Brazil 5.6%, Singapore 5.5%
Bangladesh	US 13.9%, Germany 12.9%, UK 8.9%, France 5%, Spain 4.7%	China 22.4%, India 14.1%, Singapore 5.2%
Belgium	Germany 16.9%, France 15.5%, Netherlands 11.4%, UK 8.8%, US 6%, Italy 5%	Netherlands 16.7%, Germany 12.7%, France 9.6%, US 8.7%, UK 5.1%, Ireland 4.7%, China 4.3%
Brazil	China 18.6%, US 12.7%, Argentina 6.7%, Netherlands 5.3%	China 17.9%, US 15.6%, Germany 6.1%, Argentina 6%
Canada	US 76.7%	US 53.1%, China 12.2%, Mexico 5.8%
China	US 18%, Japan 6%, South Korea 4.5%	South Korea 10.9%, US 9%, Japan 8.9%, Germany 5.5%, Australia 4.1%
Colombia	US 27.5%, Panama 7.2%, China 5.2%, Spain 4.4%, Ecuador 4%	US 28.8%, China 18.6%, Mexico 7.1%, Germany 4.2%
Ecuador	US 39.5%, Peru 5.1%, Vietnam 4.3%, Colombia 4.3%	US 27.1%, China 15.3%, Colombia 8.3%, Panama 4.9%
Egypt	Saudi Arabia 9.1%, Italy 7.5%, Turkey 5.8%, UAE 5.1%, US 5.1%, UK 4.4%, India 4.1%	China 13%, Germany 7.7%, US 5.9%, Turkey 4.5%, Russia 4.4%, Italy 4.4%, Saudi Arabia 4.1%
France	Germany 15.9%, Spain 7.3%, US 7.2%, Italy 7.1%, UK 7.1%, Belgium 6.8%	Germany 19.5%, Belgium 10.7%, Italy 7.7%, Netherlands 7.5%, Spain 6.8%, US 5.5%, China 5.4%, UK 4.3%
Germany	US 9.6%, France 8.6%, UK 7.5%, Netherlands 6.6%, China 6%, Italy 4.9%	Netherlands 13.7%, France 7.6%, China 7.3%, Belgium 6%, Italy 5.2%, US 4.7%, UK 4.2%

Continued on next page

Country	Export Partners	Import Partners
India	US 15.2%, UAE 11.4%	China 15.5%, UAE 5.5%, Saudi Arabia 5.4%, US 5.2%
Indonesia	Japan 12%, US 10.8%, China 10%, Singapore 8.4%, India 7.8%, South Korea 5.1%, Malaysia 5.1%	China 20.6%, Singapore 12.6%, Japan 9.3%, Malaysia 6%, South Korea 5.9%, Thailand 5.7%, US 5.3%
Iran	China 22.2%, India 9.9%, Turkey 8.4%, Japan 4.5%	UAE 39.6%, China 22.4%, South Korea 4.7%, Turkey 4.6%
Ireland	US 23.7%, UK 13.8%, Belgium 13.2%, Germany 6.6%, Netherlands 4.4%, France 4.4%	UK 32.5%, US 14%, France 10.2%, Germany 9.3%, Netherlands 4.9%, China 4.1%
Israel	US 27.5%, UK 6.1%, China 4.9%	US 13%, China 9.3%, Germany 6.1%, Belgium 5.3%, Italy 4%
Italy	Germany 12.3%, France 10.3%, US 8.7%, UK 5.4%, Spain 4.8%	Germany 15.4%, France 8.7%, China 7.7%, Netherlands 5.6%, Spain 5%, Belgium 4.7%
Jamaica	US 24.4%, Canada 16.5%, Russia 9.3%, Netherlands 8.9%, UK 6.5%	US 32.6%, China 12%
Japan	US 20.2%, China 17.5%, South Korea 7.1%, Thailand 4.5%	China 24.8%, US 10.5%, Australia 5.4%, South Korea 4.1%
Korea, South	China 26%, US 13.3%, Vietnam 5.3%, Japan 4.9%	China 20.7%, Japan 10.5%, US 10.1%, Germany 4.8%, Saudi Arabia 4.5%
Lebanon	Saudi Arabia 12.1%, UAE 10.6%, South Africa 6.6%	China 11.5%, Italy 7.1%, Germany 6.8%, France 6%, US 5.7%, Russia 4.6%
Malaysia	Singapore 13.9%, China 13%, Japan 9.5%, US 9.4%, Thailand 5.7%, India 4.1%	China 18.8%, Singapore 12%, US 8.1%, Japan 7.8%, Thailand 6.1%, South Korea 4.5%, Indonesia 4.5%

Continued on next page

Country	Export Partners	Import Partners
Malta	Germany 13.3%, France 10.2%, Singapore 7.3%, UK 6.4%, US 5.8%, Italy 5.6%, Japan 4.7%	Italy 23%, Netherlands 8.4%, UK 7.5%, Germany 6.8%, Canada 6.1%, China 4.1%, France 4%
Mexico	US 81.1%	US 47.3%, China 17.7%, Japan 4.4%
Morocco	Spain 22.1%, France 19.7%, India 4.9%, US 4.3%, Italy 4.3%	Spain 13.9%, France 12.4%, China 8.5%, US 6.5%, Germany 5.8%, Italy 5.5%, Russia 4.4%, Turkey 4.3%
Netherlands	Germany 24.5%, Belgium 11.1%, UK 9.3%, France 8.4%, Italy 4.2%	Germany 14.7%, China 14.5%, Belgium 8.2%, US 8.1%, UK 5.1%
Oman	China 35.4%, UAE 15.3%, South Korea 6.8%, Saudi Arabia 5.8%, Pakistan 4.2%	UAE 29.7%, Japan 10.2%, US 7.5%, China 6.7%, India 6.3%
Pakistan	US 13.1%, UAE 9.1%, China 8.8%, UK 5.4%, Germany 4.9%	China 28.1%, Saudi Arabia 10.9%, UAE 10.8%
Panama	US 19.7%, Germany 13.2%, China 5.9%, Netherlands 4.1%	US 25.9%, China 9.6%, Mexico 5.1%
Peru	China 22.1%, US 15.2%, Canada 7%	China 22.7%, US 20.7%, Brazil 5.1%, Mexico 4.5%
Philippines	Japan 21.1%, US 15%, China 10.9%, Singapore 6.2%, Germany 4.5%, South Korea 4.3%	China 16.2%, US 10.8%, Japan 9.6%, Singapore 7%, South Korea 6.5%, Thailand 6.4%, Malaysia 4.8%, Indonesia 4.4%
Puerto Rico	Canada 18.6%, Mexico 15.7%, China 7.7%, Japan 4.2%	China 21.5%, Canada 13.2%, Mexico 13.2%, Japan 5.9%, Germany 5.5%
Russia	Netherlands 11.9%, China 8.3%, Germany 7.4%, Italy 6.5%, Turkey 5.6%, Japan 4.2%	China 19.2%, Germany 11.2%, US 6.4%, Italy 4.6%

Continued on next page

Country	Export Partners	Import Partners
Saudi Arabia	China 13.2%, Japan 10.9%, US 9.6%, India 9.6%, South Korea 8.5%	China 13.9%, US 12.7%, Germany 7.1%, South Korea 6.1%, India 4.5%, Japan 4.4%, UK 4.3%
Singapore	China 13.7%, Malaysia 10.8%, Indonesia 8.2%, US 6.9%, Japan 4.4%, South Korea 4.1%	China 14.2%, US 11.2%, Malaysia 11.2%, Japan 6.3%, South Korea 6.1%, Indonesia 4.8%
South Africa	China 11.3%, US 7.3%, Germany 6%, Japan 4.7%, UK 4.3%, India 4.2%	China 17.6%, Germany 11.2%, US 6.7%, India 4.7%, Saudi Arabia 4.1%
Spain	France 15.7%, Germany 11%, Italy 7.4%, UK 7.4%, US 4.5%	Germany 14.4%, France 11.7%, China 7.1%, Italy 6.5%, Netherlands 5%, UK 4.9%
Sri Lanka	US 26%, UK 9%, India 7.2%, Germany 4.3%	India 24.6%, China 20.6%, UAE 7.2%, Singapore 5.9%, Japan 5.7%
Taiwan	China 27.1%, US 10.3%, Japan 6.4%, Singapore 4.4%	Japan 17.6%, China 16.1%, US 9.5%
Thailand	US 11.2%, China 11.1%, Japan 9.4%, Malaysia 4.8%, Australia 4.6%, Vietnam 4.2%, Singapore 4.1%	China 20.3%, Japan 15.4%, US 6.9%, Malaysia 5.9%, UAE 4%
Turkey	Germany 9.3%, UK 7.3%, Italy 4.8%, US 4.5%, France 4.1%	China 12%, Germany 10.3%, Russia 9.9%, US 5.4%, Italy 5.1%
UAE	Iran 14.5%, Japan 9.8%, India 9.2%, China 4.7%, Oman 4.3%	China 15.7%, India 12.8%, US 9.7%, Germany 6.8%, UK 4.4%
UK	US 14.6%, Germany 10.1%, China 6%, France 5.9%, Netherlands 5.8%, Ireland 5.5%	Germany 14.8%, China 9.8%, US 9.2%, Netherlands 7.5%, France 5.8%, Belgium 5%
US	Canada 18.6%, Mexico 15.7%, China 7.7%, Japan 4.2%	China 21.5%, Canada 13.2%, Mexico 13.2%, Japan 5.9%, Germany 5.5%

Continued on next page

Country	Export Partners	Import Partners
Vietnam	US 21.2%, China 13.3%, Japan 8.4%, South Korea 5.5%, Germany 4.1%	China 34.1%, South Korea 14.3%, Singapore 6.5%, Japan 6.4%, Thailand 4.5%

This data allows us now to create a matrix of flows between the single container ports. It will represent supplies and demands of each node of the network.

CHAPTER 3:
Analysis I – Network Science Analysis

Chapter 2 describes how we select and collect the data for our model. It requires decisions about the selection of nodes, which of the real world connections we want to be modeled as edges and where we want to set thresholds, in order to delimit the granularity of our network.

In this chapter, we introduce and use common network science terms and definitions. We follow the lexicon introduced in the course *MA4404: Structure and Analysis of Complex Networks* by Gera in 2017 [29] and adapted from the book *Networks: An Introduction* by M. E. J. Newman [30].

We use several software tools to support our modeling and visualization of results. Foremost among them is *Gephi* [31], an open-source tool for network analysis and visualization. We also use the plotting Python library *matplotlib* [32]. When building and manipulating the network, we use the Python library *NetworkX* (http://networkx.github.io/), which is specialized for studying large real-world graphs and networks.

3.1 Network Metrics and Properties

As described in Chapter 2, our network consists of two separate sub-networks – the sea layer and the road layer. Both layers evolve in parallel and operate on the same set of nodes, so we refer to the network as a *multiplex network* [33]. The edges of the network are undirected, so they can be traversed in both directions.

3.1.1 Sea Layer

The sea layer is the main layer of our network and the primary focus of our study. It consists of 120 nodes, that represent 94 container ports and 26 maritime chokepoints. The edges of the layer represent the maritime ship routes between those container ports and/or chokepoints. The layer is structured such that there exists a path between each pair of nodes, which means that there are no disconnected components. There are a total of 1,518 edges

in the sea layer. Figure 3.1 presents a visualization of the sea layer. The red points are the container ports and the yellow points represent the maritime chokepoints. The lines represent the direct connections in between, visualized here simply as straight lines.

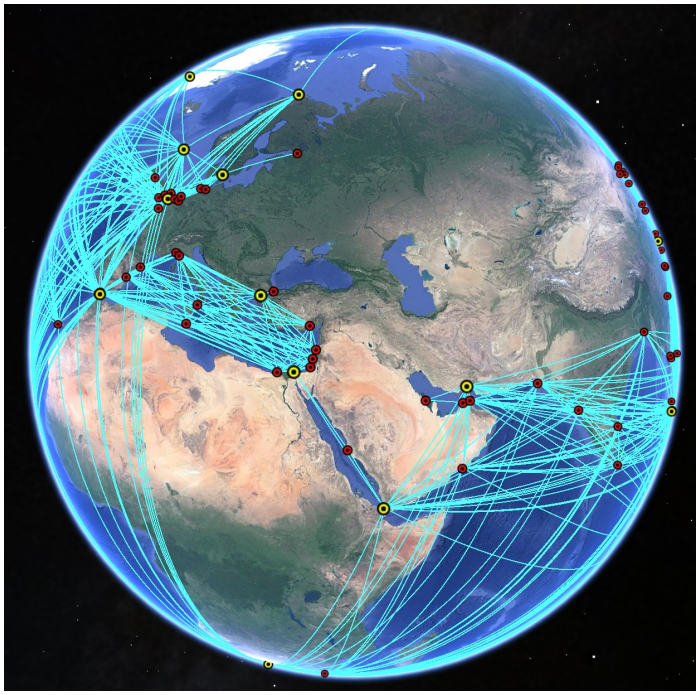


Figure 3.1. The sea layer visualized in Google Earth.

The structure of the sea layer is displayed in Figure 3.2. It shows the geo-references nodes and the layout corresponds to a world map. The reader may identify the seaports of North and South America on the left side of the graph, Europe and Africa in the middle, a white space for the Asian continent in the top right corner and Australian ports on the bottom right corner. Easily recognizable is the Mediterranean Sea in the middle of the graph. The edges appear as straight lines, but of course, the distances between the corresponding edges have been calculated for the real world routes. For better visibility of the graph, the edges that represent the connections across the Pacific Ocean have been omitted from the graph.

The natural spatial pattern of the seaport locations around the world defines the structure of the sea layer. The seaports are always situated at one specific body of water, like the Atlantic Ocean or the South China Sea. By construction, the maritime chokepoints always separate the single bodies of water from their neighbors. The outcome of this is that the container ports are subdivided into eleven groups, corresponding to the adjacent bodies of

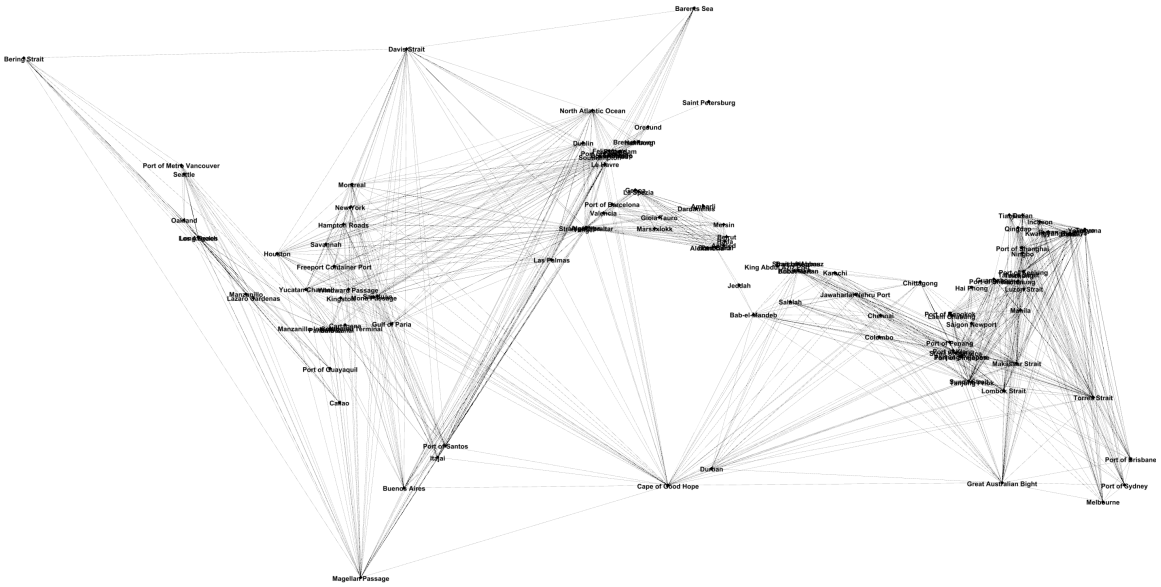


Figure 3.2. Sea routes network graph.

water. These are in particular:

- Pacific Ocean
- Atlantic Ocean
- Indian Ocean
- Caribbean Sea
- North Sea
- Baltic Sea
- Mediterranean Sea
- Sea of Marmara
- Red Sea
- Persian Gulf
- South China Sea.

Because the maritime chokepoints in the graph act as transition nodes for longer routes, we only need to model the connections between ports of each body of water separately. Because of the existence of transition points, connections between ports of different bodies of water are not required. But within each body of water, all possible pairwise connections are implemented in the model, which makes each subgraph of each body of water a clique.

The chokepoint nodes always count as members of both neighboring cliques.

Definition 1. *A clique is a maximum complete subgraph in which all nodes are adjacent to each other [30].*

Definition 2. *A k-clique is a clique of k nodes [30].*

Because each body of water is a complete graph, the sea layer as a whole is a collection of eleven cliques, each interconnected to others by one or more chokepoints. The smallest cliques are 2-cliques, which represent the Baltic Sea, the Sea of Marmara or the Persian Gulf. The biggest clique is a 37-clique and it represents the Pacific Ocean. This structure makes this layer a clustered graph, so its adjacency matrix can be written in a block form. Figure 3.3 shows this adjacency matrix, where the cluster blocks can be easily recognized in the white and grey areas. Matrix cells with entries stand for existing connections between the corresponding nodes. The blue area shows the connections to, from and between the chokepoints. There is no obvious distinguishable structure. The clustering coefficient of the sea layer is very high with a value of 0.9, where a value of 1 is a maximal value. This is due to the clique structure of the subgraph.

Definition 3. *The clustering coefficients measure the average probability that two neighbors of a vertex are themselves neighbors (a measure of the density of triangles in a network) [30].*

Definition 4. *A community in a network is a subset of nodes that share common or similar characteristics, based on which they are grouped [30].*

According to the classification of Jeub et al. [34] for block models of network adjacency matrices, the adjacency matrix (at least for the white and grey areas) of our sea layer has a *low-dimensional structure* because of high spectral clustering. This is because the communities of the layer are the cliques of nodes with the separate bodies of water, and the only overlapping communities in the layer are chokepoints, which always belong to two or three communities.

Based on the structure classification of Clauset for network communities [35], the sea layer has an *assortative community* structure. This again means that the nodes with similar characteristics (here, the membership in a body of water) have a tendency to connect with

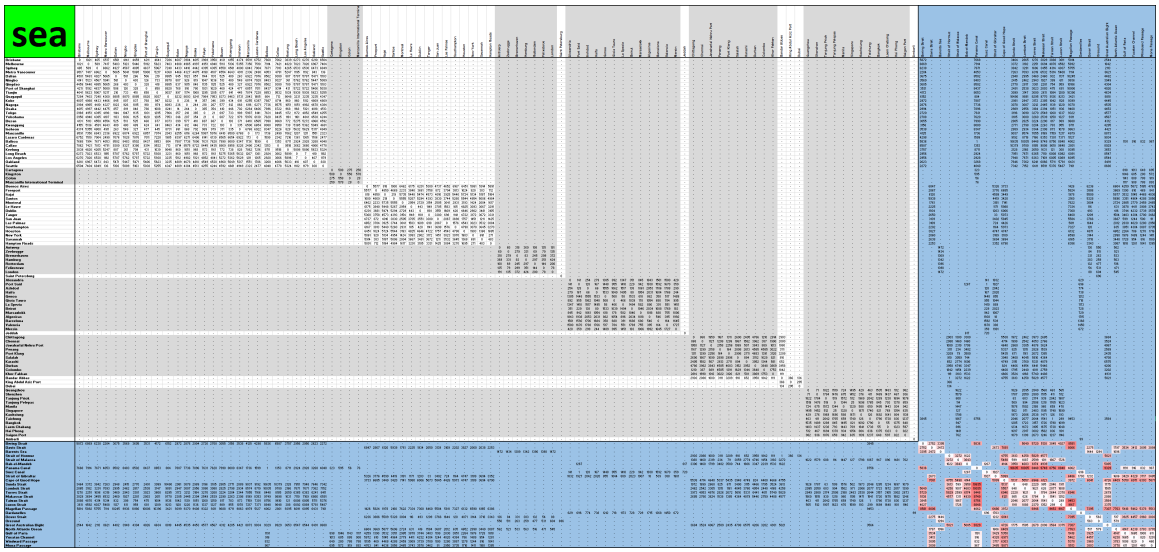


Figure 3.3. Adjacency matrix of the sea layer.

each other, which we often expect for a "community structure." There are no entries outside of the white and grey areas in the adjacency matrix in Figure 3.3 and therefore, we do not see any evidence of a core-periphery.

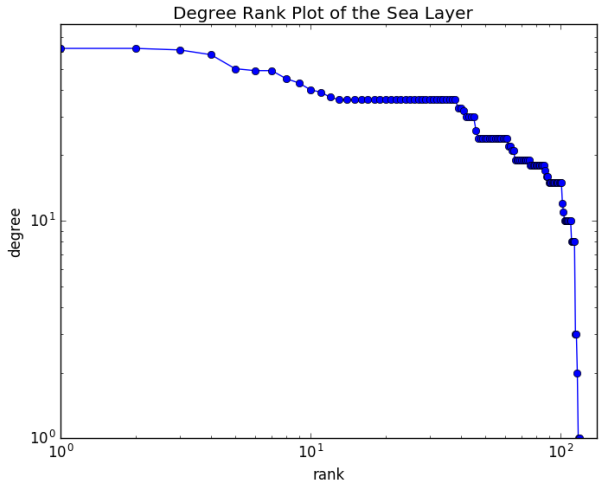


Figure 3.4. Degree distribution of the sea layer.

Figure 3.4 displays the degree distribution of the sea layer. The maximum degree is 62, which is the node degree of multiple maritime chokepoints connecting the Pacific Ocean, the Indian Ocean and the South China Sea, each of which contains many ports. The minimum

degree is one, corresponding to the ports in bodies of water with only one outlet, like Saint Petersburg or Ambarli. The average degree of the sea layer is 25.3. The overall average shortest path length is 2.2, which means that travelling between two ports, ships have to pass on average 1.2 chokepoints. The longest shortest path within the layer has a length of five.

3.1.2 Road Layer

The road layer plays a subordinate role in our network model. It is implemented to provide a possibility of alternate routes for goods transported by sea. Whenever there is a situation where the destination container port is out of order or a maritime chokepoint along the route is impassable, a change to landside transportation of the road can be considered. More specifically, road transportation can be used either to the destination port or to the next available container port for further transportation by sea.

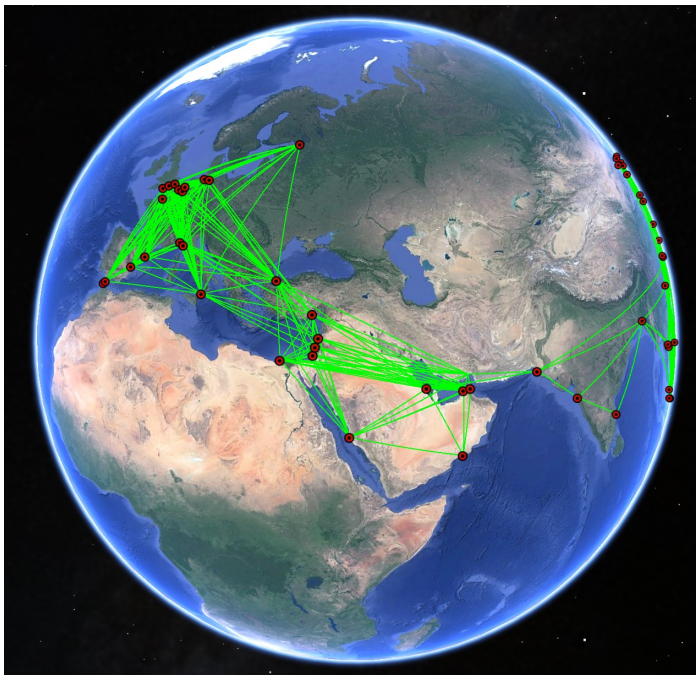


Figure 3.5. The road layer visualized in Google Earth.

Because our network is a multiplex network, we utilize the same nodes for the road layer as we do for the sea layer. However, as the maritime chokepoints only have a meaning for sea routes, we omit these transition nodes in the road layer. Since it is theoretically possible to

travel by car between France, South Korea and South Africa without using any ferries, we impose a threshold for road connections at maximum 2,000nmi, as explained in Chapter 2. As a result, some nodes that are too far away from others, like Durban or Tangier, are now completely disconnected from others. We omit these nodes from the subgraph, since a transition to landside transportation there is senseless. This results in a well-arranged road layer, as visualized in Figure 3.5.

The resulting road layer contains 83 nodes and 365 edges. Figure 3.6 provides a global overview of the layer. The reader may recognize North and South America on the left side with four connected components. Furthermore, there is still a large component connecting Europe, Northern Africa and Asia. Three other components are in Taiwan, Japan and Australia, resulting in a total of eight connected components.

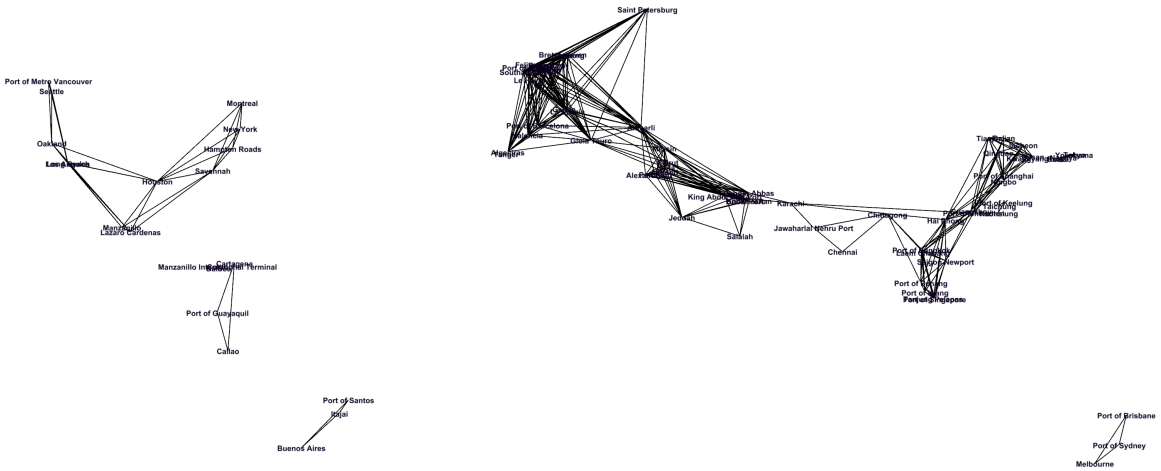


Figure 3.6. Road routes network graph.

The average degree of this layer is approximately 8.8. The minimum is 2, which applies to all five small components, each containing only three nodes, like in Australia or South America. The maximum degree is 24, corresponding to Ambarli, Turkey; this node has a central location within its component, being connected to all European ports and some of the Asian ports. Interestingly, Ambarli has the smallest degree of all nodes in the sea layer. The degree distribution of the road layer is shown in Figure 3.7. The overall average shortest path length is 3, but here we observe the components for North America and Eurasia separately. The Eurasian component has an average shortest path length of 3.1, and in North America it is 1.6. For all other components, it is 1, of course.

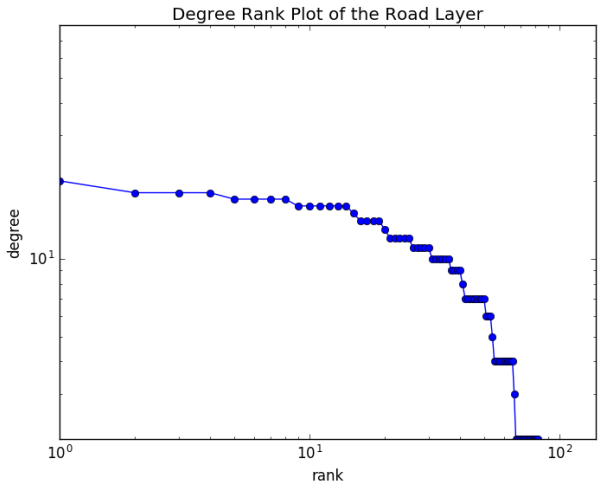


Figure 3.7. Degree distribution of the road layer.

The overall clustering coefficient of the road layer is 0.87, which is still high. This is because of the five small components, each with a value of 1. The values for North America and Eurasia are 0.82 and 0.83, respectively. The high values follow from the fact of the selected threshold to introduce distance based edges to our model. Because of that, mostly the nodes with a similar degree are connected, like it is in the case in Europe, and there is no connection to the weakly connected Middle East.

3.2 Centralities

In this subsection, we try to answer the question, which nodes are "important" in our network. For this, we consider an analysis of vertex centrality, which originated from the social networks research.

Definition 5. *Centralities quantify the intuitive notion of importance of a node(s) in a network [30].*

There a many different types of measures of a centrality of vertices in a network. We analyze only the degree centrality, closeness centrality and betweenness centrality.

The *degree centrality* is based on the degree of each node. It expresses the number of vertices that a node i influences directly and is a metric of a local effect [30]. It doesn't

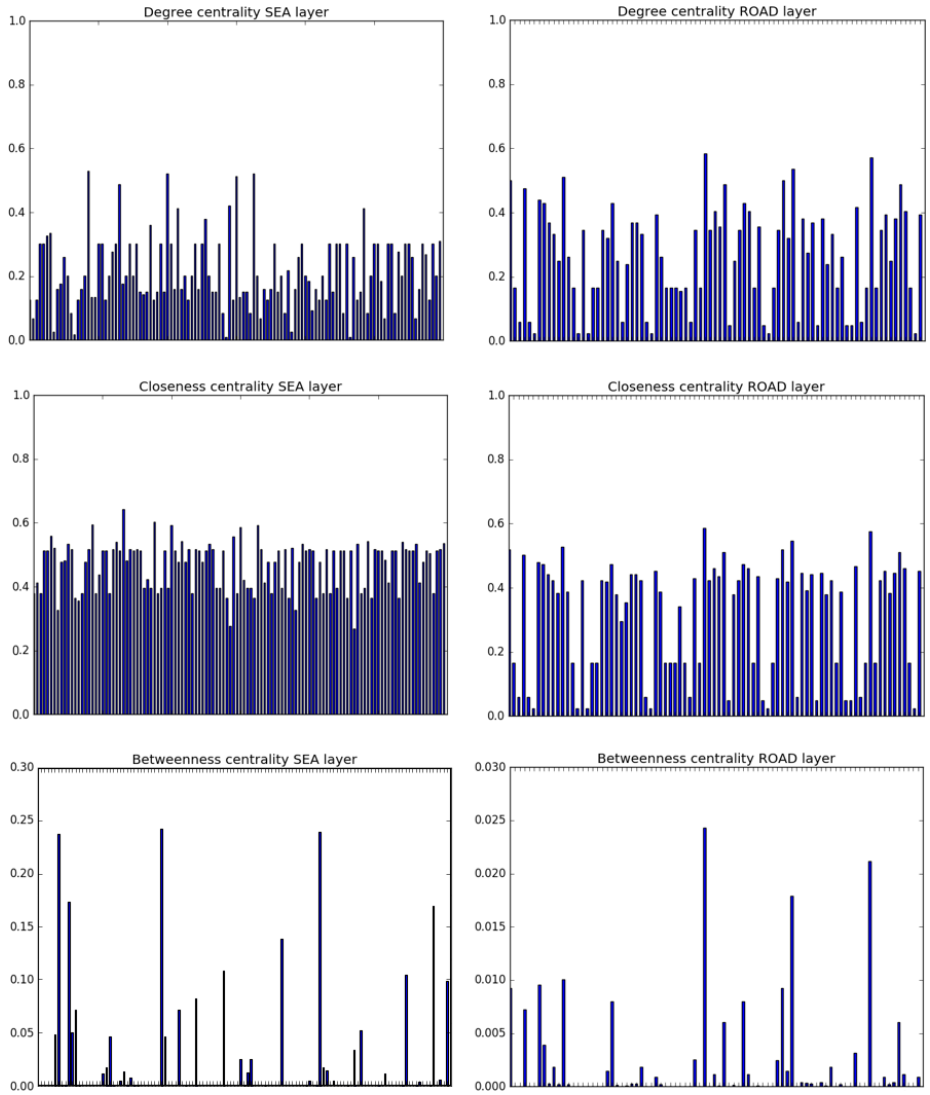


Figure 3.8. Plots of centralities' distributions of the sea and the road layer. The y-axes represent the centrality values on a scale from zero to one. The x-axes represent the network nodes in an arbitrary order, which is consistent across all plots.

incorporate any non-local information in the network. Degree centrality tend to be used as a measure if the number of one-hop connections from and to the node is important.

The distributions for the centralities of both layers appear in Figure 3.8. The sea layer nodes with the highest degree centrality have a value of about 0.53; these correspond to the transition nodes between the Pacific Ocean, the Indian Ocean and the South China Sea, that

have the highest node degree, as introduced in Subsection 3.1.1. Respectively, the nodes with the lowest degree have the lowest degree centrality of about 0.01. A similar situation can be observed within the road layer, where Ambarli has the highest centrality at about 0.29 and the nodes of the small connected components have the lowest value of about 0.02.

Newman [30] describes the *closeness centrality* as follows:

> Closeness is based on the average distance between a vertex i and all vertices in the graph. Closeness centrality can be viewed as the efficiency of a vertex in spreading information to all other vertices.

Thereby, the number of direct "neighbors" of a node i is less important for the closeness centrality, but rather the number of nodes with a short distance from i. The higher the centrality is, the closer is i to all others. We apply the normalized closeness centrality on a scale from zero to one, so we can compare the values between both layers.

Within the sea layer, the nodes have more or less significantly higher values for the closeness centrality, than for the degree centrality. But an interesting fact is that the five nodes with the highest centrality are the same as they were for the degree centrality. This is also the case for the road layer. For the other nodes as well, the ranking is more or less similar with only few changes. For the closeness centrality, the difference between the highest value of 0.64 and the lowest value of 0.27 is lower than for the degree centrality. For the road layer, the centrality can only be computed per component, there are no shortest paths across the components. There, the ranking stays similar as well. But in general the values don't increase, and the plot looks similar to the degree centrality.

The *betweenness centrality* has a high value, if a node i is a part of many shortest paths between other nodes [30]. This measure can be used if the flow in a network is analyzed, for example, a flow of data packages or goods. Due to the influence and importance of node with a high betweenness centrality, its removal can drastically disrupt the flow of a network.

To calculate the betweenness centralities in our network, we evaluate it as a weighted graph. The weights of the edges are the real sea distances between the nodes. In general, most values of the betweenness centrality are much lower, than for the other centralities. In both layers, many nodes have a value of zero, which means, that their geographical location does

not qualify them as a hub on routes between other nodes. Within the sea layer, the Strait of Gibraltar, the Suez Canal and Bab-el-Mandeb are by far the most "central" nodes with values of 0.217, 0.215 and 0.212, respectively. These are the chokepoints on the shortest route between Asia and Europe. They are followed by the Strait of Malacca with a value of 0.115. Surprisingly, the Panama Canal has a value of only 0.11. Of course, its importance is based on its throughput, which is not taken into account here. The most "central" nodes within the road layer are Karachi, Shenzhen and Ambarli with values of 0.19, 0.12 and 0.1, respectively. This is not surprising, since they act as hubs on many road routes in Asia and between Asia and Europe.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 4:
Analysis II – Network Flows / Resilience

Chapter 3 identifies nodes that are central and therefore potentially important to the network. In this chapter, we implement a linear optimization model that also includes network capacity and throughput, and we use it to analyze the cargo flows within the global maritime transportation network.

Our analysis begins with an implementation of a "base case," in which all container ports and maritime chokepoints are operable. In the next step, we consider disruptions to one or more ports and/or chokepoints and observe the ability of the system to adapt to the breakdown. In this manner, by considering the potential consequences associated with a variety of "what if" scenarios, we seek to identify the nodes that are critical to the global flow of cargo traffic. Finally, we compare our findings with the previously identified central nodes to contrast these two perspectives on what it means for a node to be critical.

4.1 Linear Program Formulation

To represent the global cargo container flow, we implement a linear optimization model that utilizes not only connectivity data but also distances, capacities and demands, as described in Chapter 2. The cargo moves along the network edges from node to node to get from its origin to its destination. Every port of the network either exports cargo to trade partner ports or imports it from other ports or it does both. To adjust our existing network to model the situation properly, in the first step we replace each undirected edge by two directed arcs, one in each direction. Furthermore, in order to be able to close ports to simulate breakdowns within the model, we apply the node splitting method, where we create an additional node for each existing one. An example is shown in Figure 4.1.

The example in Figure 4.1 shows on the left the situation prior to node splitting, with undirected edges between the nodes. It shows two container ports that have a direct sea connection in between (the blue edge), a connection via a maritime chokepoint and a road connection (green edges). The figure on the right depicts the same situation after node splitting: each edge is replaced by two bidirectional arcs with the exception of chokepoints.

These are now split in an incoming and an outgoing node with exactly one arc in between. If we want to disable a chokepoint, we target the arc from the incoming to the outgoing node. To disable a port, we have to do the same on both arcs between the seaward and the landward nodes of the affected port.

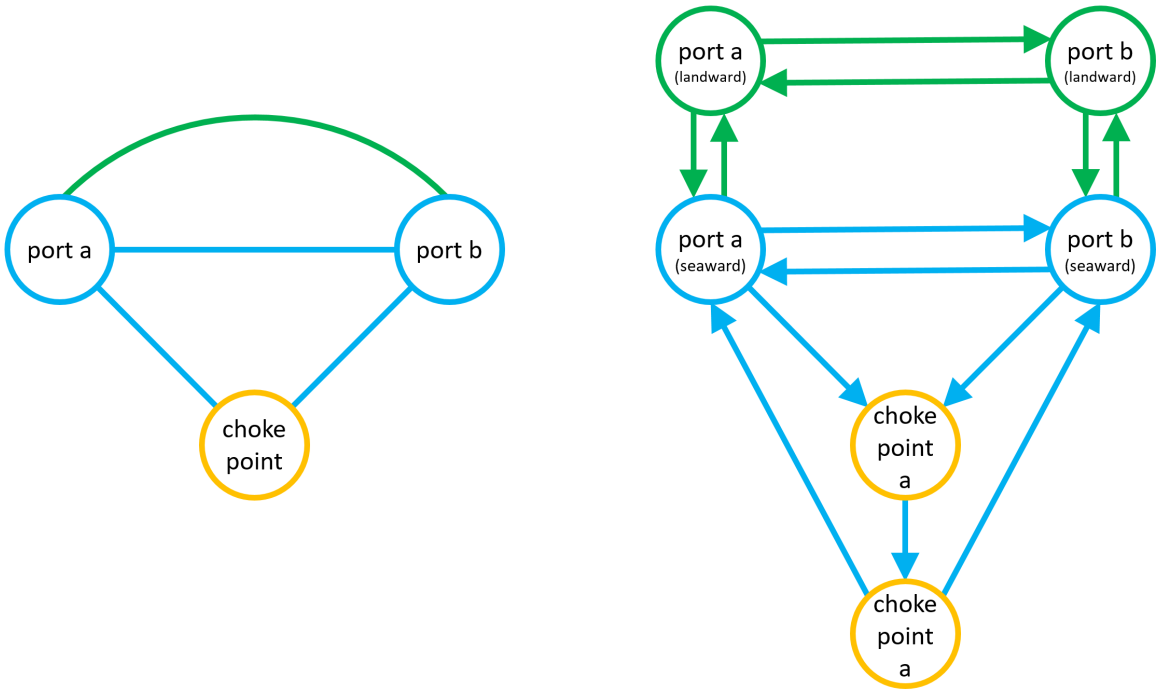


Figure 4.1. Application of node splitting. Left figure shows a network, and the right one shows the same network after node splitting. The blue lines are sea connections and green lines are road connections.

One way to implement a disruption on an arc could be remove it from the network, or perhaps reduce its capacity to zero (which would have the same effect). However, this has the potential to create infeasible network flows and can be problematic when solving iteratively for solutions. Instead, we use "cost-based interdiction" and increase the usage cost of each targeted arc as described in [36]. In this way, a targeted arc becomes too expensive and will not be considered for a solution.

Container ports and canals in the real world can only process a finite number of ships in a certain time period. This number is referred to as their *capacity*. Accordingly, we implement the node capacities in our network: since our cargo flows are measured in TEUs, we apply these units to the capacities as well. A study of the Tioga Group, Inc. [37] provides

a broad overview over port capacities of the major U.S. ports, subdivided in North Atlantic, South Atlantic and Gulf Coast regions. The container terminal capacity is measured in five "dimensions" such as berth length and depth, number of berths and cranes, container yard acreage and the operating hours of the ports (number of shifts). The study comes to a conclusion that based on the five "dimensions", ports have different capacities and the distribution of strengths and weaknesses over the "dimensions" is different in each port. But in general, each of the three regions on its own could handle roughly double of the actual throughput before reaching its capacity constraints. Based on this study, we assume that we can transfer this behavior to all ports of our network. Therefore, we assign the double throughput of each container port as its maximal capacity.

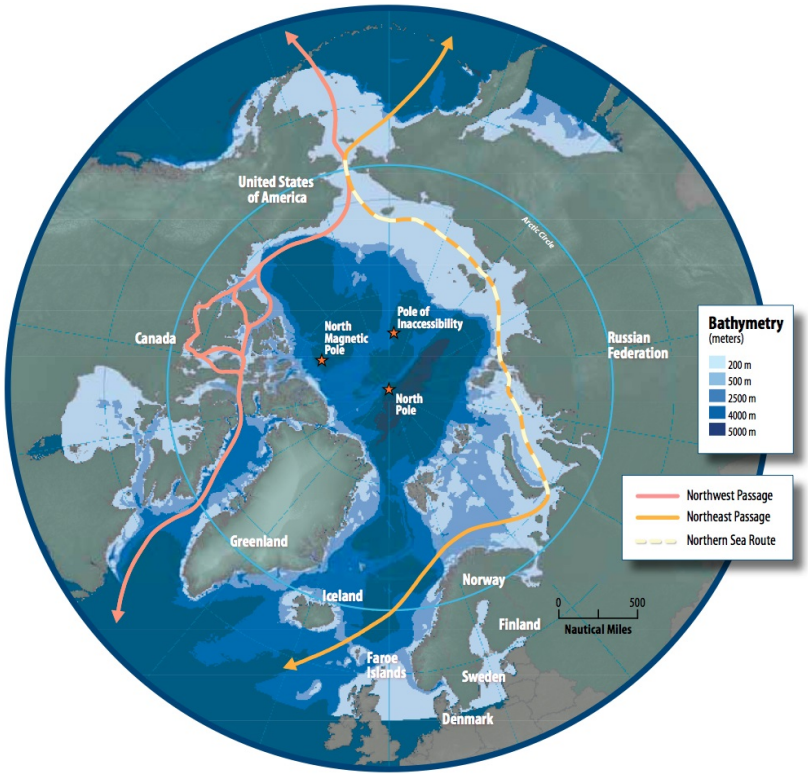


Figure 4.2. Arctic sea routes. Source: [38]

For the maritime chokepoints, we will implement capacities for only two nodes. The first is the Panama Canal, which is limited by the operation of the locking system. According to the Panama Canal Authority, the canal's capacity in 2012 was about 285,000,000 tons per year [39]. Assuming an average of fourteen tons per TEU, we set the capacity for the Panama Canal in out model to 20,357,143 TEUs. The second chokepoint where we

implement a maximum capacity is the Bering Strait. Because of the arctic ice melting, maritime transportation through the arctic sea routes, as shown in Figure 4.2, becomes more likely. But during the most months of the year, the routes are not free from ice. Therefore, to set a realistic capacity the Arctic routes, we take the total number of complete through transits in 2016, which is eighteen, according Murray [40]. Assuming the average of 3,649 TEUs per container ship, according to a United Nations Conference on Trade and Development (UNCTAD) report [41], set implement the capacity of 65,682 TEUs in the Bering Strait.

In terms of our estimate for demand for transport of container cargo worldwide, Subsection 2.2.2 presents an approximation of the cargo flow between each pair of ports for our model. We build this estimate based on the throughput in each port (Table 2.1), the export and import volumes of each country (Table 2.3) (in particular their ratio), and finally the information about the export and import partners and their proportion of trade for each country (Table 2.4). For each pair of ports, we treat the resulting volume of flow as a demand in the destination port for the cargo from the port of origin. Table 4.1 shows such an example, where each port has multiple demands for cargo from others. Not every port has demands for cargo from every other port. Some pairs of ports have only a unidirectional flow or no trade flow at all. To ensure that the required demands of other ports are provided by each port, the main diagonal contains the sum of all demands from the corresponding port as its supply value. This ensures balance between supplies and demands within the model. Because all supplies and demands occur on land, we assign the cargo demands to the (landward) road nodes instead of their corresponding (seaward) sea nodes.

Table 4.1. Example of cargo demands (in TEU).

		port of origin			
		Hamburg	Rotterdam	Barcelona	Gioia Tauro
port of destination	Hamburg	-450,000	70,000	110,000	160,000
	Rotterdam	80,000	-400,000	140,000	0
	Barcelona	200,000	230,000	-250,000	0
	Gioia Tauro	170,000	100,000	0	-160,000

Although we only deal with containerized cargo in our model, we actually have different types of commodities, since every container has its origin and destination. We next provide the mathematical formulation of the multi-commodity linear optimization model that min-

imizes the total cost of the global cargo flow. It preserves the balance of flow at the nodes, and allows for the interdiction of individual arcs.

Indices and Sets

$n \in N$	nodes (alias i, j)
$s \in S \subset N$	sea nodes
$r \in R \subset N$	road nodes
	$N = S \cup R$; $S \cap R = \varnothing$
$(i, j) \in A$	directed arc from node i to node j

Data [units]

c_{ij}	per unit cost of traversing arc $(i, j) \in A$ [dollars/TEU]
u_{ij}	upper bound on total directed flow on arc $(i, j) \in A$ [TEUs]
$\hat{x}_{ij}$	1 if arc $(i, j) \in A$ interdicted, 0 otherwise [binary]
q_{ij}	per unit penalty cost of traversing interdicted arc $(i, j) \in A$ [dollars/TEU]
d_n^r	demand at node $n \in N$ for cargo originating from node $r \in R$ [TEUs]
	(supply if $d_n^r < 0$)
p_n	per unit penalty cost for demand shortfall at node $n \in N$ [dollars/TEU]

Decision Variables [units]

Y_{ij}^r	flow on arc $(i, j) \in A$ of cargo originating from node $r \in R$ [TEUs]
Z_n^r	shortfall of cargo originating from node $r \in R$ at node $n \in N$ [TEUs]
E_n^r	excess of cargo originating from node $r \in R$ at node $n \in N$ [TEUs]

Formulation

$$\min_{Y,Z} \quad \sum_{r\in R}\sum_{(i,j)\in A}\left[\left(c_{ij}+q_{ij}\hat{x}_{ij}\right)Y_{ij}^{r}\right]+\sum_{r\in R}\sum_{n\in N}\left[p_n\left(Z_n^r+E_n^r\right)\right] \tag{D0}$$

$$\text{s.t.} \quad \sum_{(i,n)\in A} Y_{in}^{r}-\sum_{(n,j)\in A} Y_{nj}^{r}+Z_n^r-E_n^r=d_n^r \qquad \forall n\in N,\forall r\in R \tag{D1}$$

$$0\le\sum_{r\in R} Y_{ij}^{r}\le u_{ij} \qquad \forall (i,j)\in A \tag{D2}$$

$$Y_{ij}^{r}\ge 0 \qquad \forall (i,j)\in A,\forall r\in R \tag{D3}$$

$$Z_n^r\ge 0 \qquad \forall n\in N,\forall r\in R \tag{D4}$$

$$E_n^r\ge 0 \qquad \forall n\in N,\forall r\in R \tag{D5}$$

Discussion

The objective function (D0) is a summation of costs that are generated by the model: the combined flow cost over sea and road connections and the penalty cost for having a demand shortfall or excess at some nodes. The constraint (D1) ensures the balance of flow at each node and for each commodity, setting the incoming and the outgoing flow, the shortfall and the excess equal to the demand at that node. Constraints (D2), (D3), (D4) and (D5) define bounds on decision variables. The u_{ij} parameter defines a capacity for arcs, if applicable. As mentioned earlier, interdictions in the model are instantiated by making the corresponding arcs very expensive, which would clearly increase the objective function value.

4.2 Preliminary example network

To gain an understanding of the optimization model and verify its correct functioning, at least by visual inspection, we test it first on a small and manageable network. For this purpose, we extract a subset of our global maritime transportation network and utilize it as a test network for initial computation and analysis. The chosen subset is located in Europe and contains three maritime chokepoints Dover Strait, Strait of Gibraltar and the strait north of Great Britain, between the Scottish Mainland and the Orkney Islands (referred below unofficially as Scottish Strait). Furthermore, it contains container ports from France, Germany, Great Britain, Italy and Malta, which are in particular:

- Bremerhaven
- Felixstowe
- Genoa
- Gioia Tauro
- Hamburg
- La Spezia
- Le Havre
- London
- Marsaxlokk
- Southampton.

After applying our node splitting convention, the test network results in 26 nodes. It contains a total of 163 directed arcs for sea and road connections. Since it is a subset of the global network, we adopt the corresponding values for distances and demands to the test network. We have ten commodities of cargo, one originating in each port. In total, there are 29 pairs of ports, that have a demand for unidirectional or bidirectional cargo exchange. The total amount of cargo in the test network is 4,096,816 TEUs. According to Zeihan [42], "modern container ships can transport goods for about net 17 cents per container mile, compared to semi-trailer trucks that do it for net $2.40, including the cost of the locomotion mode as well as operating costs in both instances." These values are valid for the U.S., but we assume the same costs for our entire network and multiply all corresponding arcs with these transportation cost values.

To set a reference for our analysis, we start with a base case without any interdictions, where the algorithm chooses the cheapest path for the cargo to its destination. We implement the linear model with *Pyomo* [43], [44], [45], a Python-based, open-source optimization modeling language, and use *IBM ILOG CPLEX Optimizer* [46] to solve it.

For the base case, the linear program solves quickly (without a noticeable solution time). Its visualization is presented in Figure 4.3. The maritime chokepoints are shown in yellow, the container ports in green. The blue lines represent cargo transportation along the corresponding real sea routes. If there was road transportation, the corresponding arcs would appear as green lines. We use straight lines for the purpose of a simplified visualization, and they should not be understood as the measured distances. The thickness of the lines is

proportional to the sum of the total amount of cargo transported on both opposing arcs.

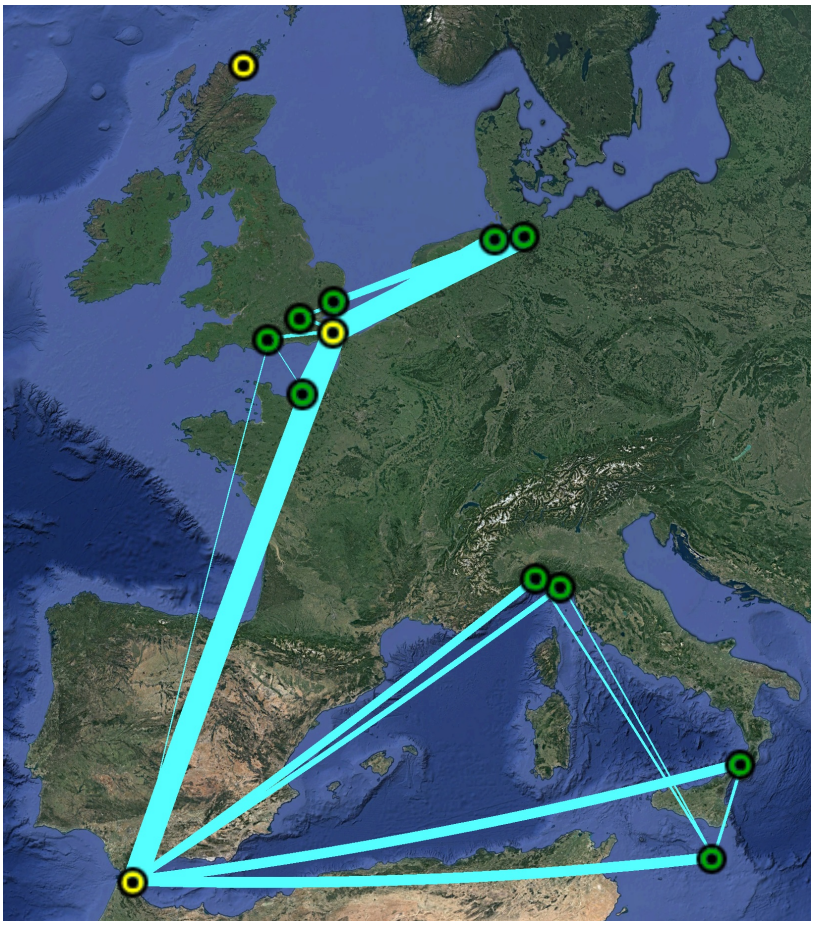


Figure 4.3. Cargo flow in the base case of the test network.

Table 4.2 shows the results of the base case scenario of the test network. The total transportation cost includes sea and road transportation. Although the mathematical formulation defines penalty cost for demand shortfall or excess at nodes, we are only interested in transportation cost. If shortfall or excess would appear in the optimal solution, we are only interested in its amount in TEUs. Due to no interdictions in the scenario, there is no shortfall, excess or road transportation. The amount of flow between chokepoints is 25% of the total cargo flow.

To stress the network, we apply one interdiction and observe its impact. We choose the port of Gioia Tauro (Italy) as the target of interdiction. Loading and offloading in the port is not possible anymore, but the cargo can still be delivered by road to the port area.

Table 4.2. Results of the base case in the test network.

Scenario Results Table	
total transportation cost	$892,301,167
total shortfall/excess	0 TEUs
portion of road transportation	0%
portion of transportation between chokepoints	25%
ports with exhausted capacity	-

Figure 4.4 shows a comparison between the base case and the interdicted port situation. The interdicted port is marked in red. It is clearly visible that there are no sea connection arcs between the interdicted port and other nodes, but additional road arcs to the port appear instead. All the cargo from Gioia Tauro is transferred first to and from the closest port of La Spezia in North Italy, which is reachable by land. But once the capacity in La Spezia is exhausted, the remaining cargo is sent through the port of Genoa (North Italy as well).

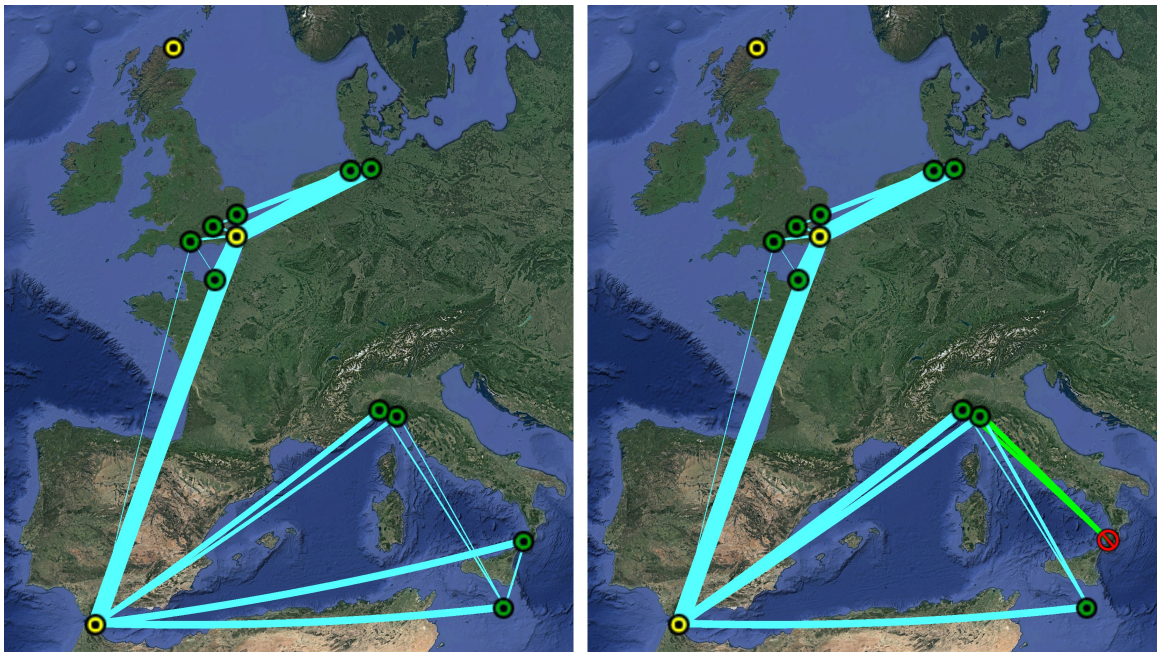


Figure 4.4. Cargo flow with an interdiction of the port of Gioia Tauro. (Base case of the left, interdicted network on the right.)

Table 4.3 shows clearly the increased costs of the total transportation, since cargo flow on the road, which is at 7% now, is much more expensive. The rerouting of cargo via close by ports prevents shortfalls in other ports. Their capacities can handle the additional load.

Table 4.3. Results of the interdiction in Gioia Tauro.

Scenario Results Table	
total transportation cost	$1,862,737,486
total shortfall/excess	0 TEUs
portion of road transportation	7%
portion of transportation between chokepoints	25%
ports with exhausted capacity	La Spezia

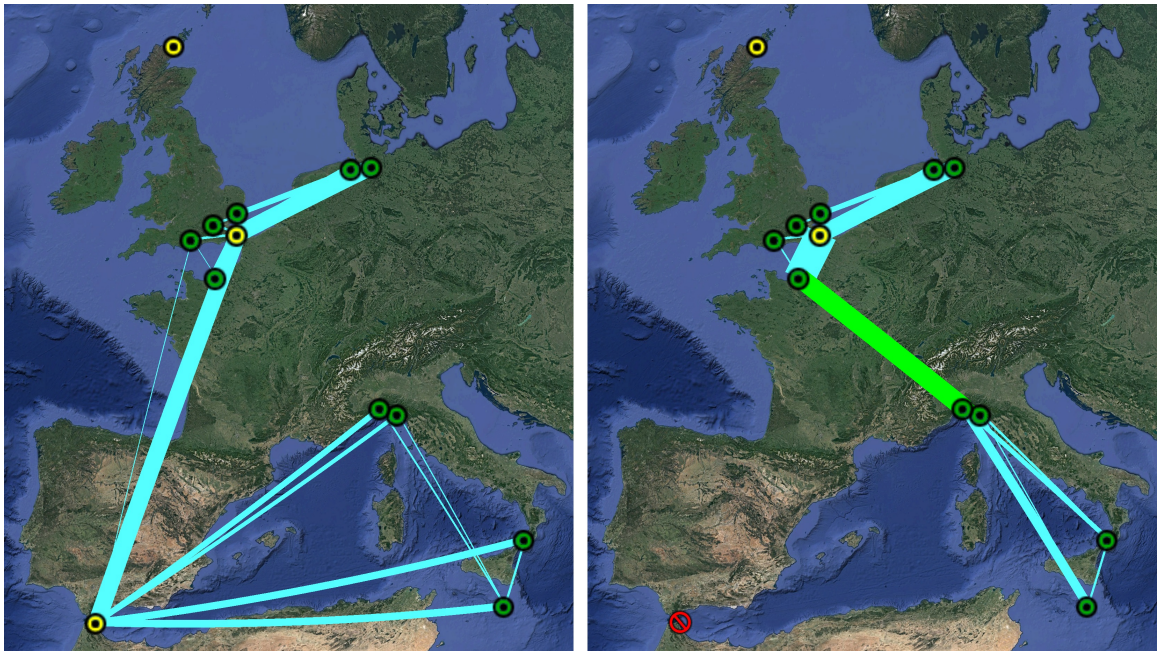


Figure 4.5. Cargo flow with an interdiction in the Strait of Gibraltar. (Base case of the left, interdicted network on the right.)

In addition, we want to observe a scenario with an interdiction on a maritime chokepoint. For this purpose, we choose the Strait of Gibraltar. It is the only sea connection between the Mediterranean Sea and the ports in the north of the test network. Figure 4.5 shows the big impact on the flow in case of the interdiction. With this chokepoint interdicted, all cargo between the two groups of ports need to be transported by road (the green edge). The ports of Genoa and La Spezia are now the hubs for the connection to Le Havre. The capacity of both ports exhausts right as the last cargo is transferred. If there was any additional cargo, the third Italian port would be needed to handle it.

As Table 4.4 indicates, the total cost is more than twice as high for the base case, since the

Table 4.4. Results of the interdiction in the Strait of Gibraltar.

Scenario Results Table	
total transportation cost	$2,973,650,229
total shortfall/excess	0 TEUs
portion of road transportation	18%
portion of transportation between chokepoints	0%
ports with exhausted capacity	Genoa, La Spezia

road transportation portion has increased to 18%. Nevertheless, still no demand shortfall of excess has to be taken by the test network.

4.3 Global Network

Our results for the preliminary example network in Section 4.2 provide confidence that the model is working properly. Therefore, we turn our attention to analyze the global maritime transportation network. As before, we start with the base case without any interdictions to the nodes. The total amount of cargo in the global network is 138,203,566 TEUs. The computation time for solving the base case on a personal laptop is approximately three seconds. But with addition of loading the data from different files and setting up the model, the computation runs for about 1.5 minutes. Generating output takes another 30 seconds, so the total computation time for one scenario is about two minutes.

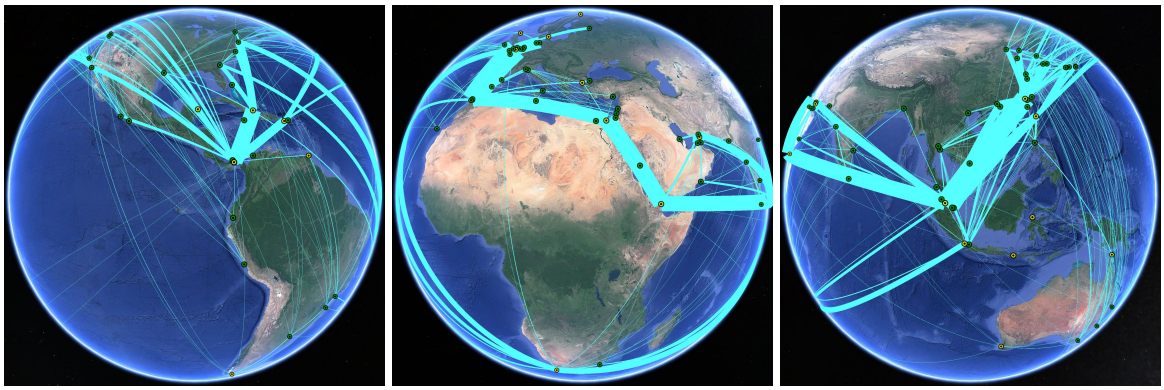


Figure 4.6. Cargo flow in the base case of the global transportation network.

The flow of the base case scenario is shown in Figure 4.6, viewed from different perspectives. It reveals the main routes of cargo transportation, distinguishable by the thick arcs around the globe. The most highly frequented routes are:

- between North America and Europe across the Atlantic,
- between Northern Europe and Asia across the Strait of Gibraltar, Suez Canal, Indian Ocean and the Strait of Malacca,
- between North America and Asia across the Atlantic, Cape of Good Hope and the Indian Ocean and
- between North America and Asia across the Pacific Ocean.

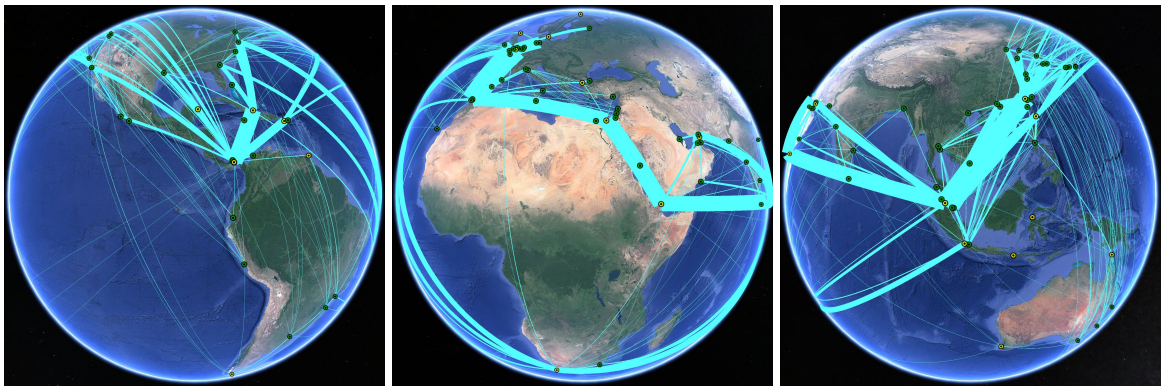
Table 4.5. Results of the base case scenario.

Scenario Results Table	
total transportation cost	$93,584,947,771
total shortfall/excess	0 TEUs
portion of road transportation	0%
portion of transportation between chokepoints	42%
ports with exhausted capacity	-

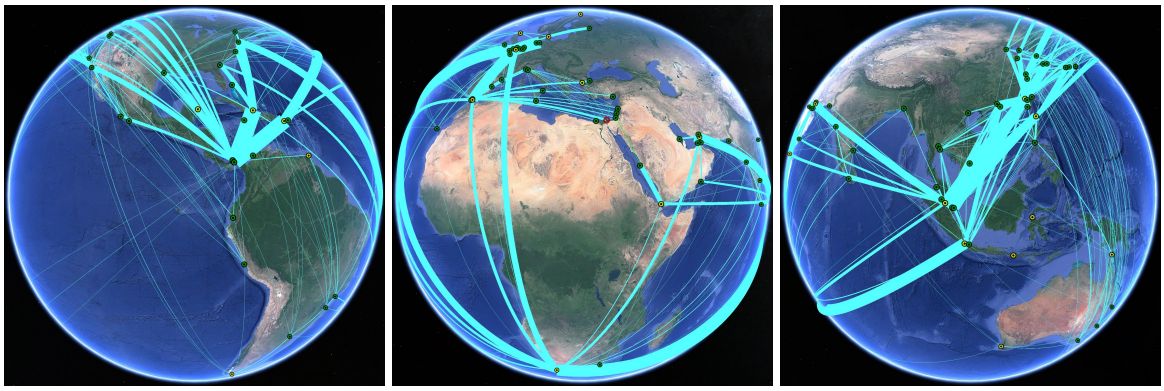
Many minor connections exist in the network as well and complete the picture of the flow. Table 4.5 presents the results of the base case. The total transportation cost sets the reference value for future scenarios with interdictions, as the lowest possible cost. Approximately 42% of the total flow is simply the transit between maritime chokepoints and we observe there is neither road transportation, nor shortfall of exhausted capacity nodes in the base case.

4.3.1 Single Interdictions

To analyze the network's behaviour under node failure, we consider an interdiction of the Suez Canal as the first scenario. Figure 4.7 illustrates the resulting cargo flow (bottom), compared to the base case cargo flow (top). The interdicted Suez Canal is highlighted by the red symbol in the lower middle figure. On the left, a slight increase of the flow between Asia and the Panama Canal can be detected, as well as an increased flow between North America and Europe. In the middle, where the Suez Canal is located, are the most significant changes. The flow between Europe and Asia across the Red Sea is completely interrupted and instead takes place over the Cape of Good Hope. The same applies for the flow between North America and Asia. On the right, we observe two effects. First, most of the cargo from Asia is directed to the Cape of Good Hope, instead of Bab-el-Mandeb (Mandeb Strait). Second, a greater portion of this flow traverses the Sunda Strait, instead



(a) Base case situation.



(b) Failure of the Suez Canal.

Figure 4.7. Comparison of the base case and the Suez Canal failure scenario.

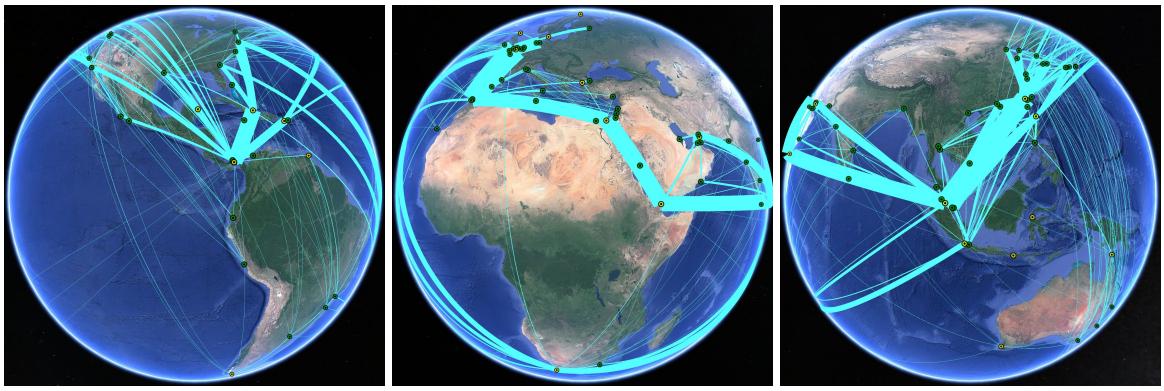
of the Strait of Malacca.

Table 4.6. Results of the interdiction of the Suez Canal.

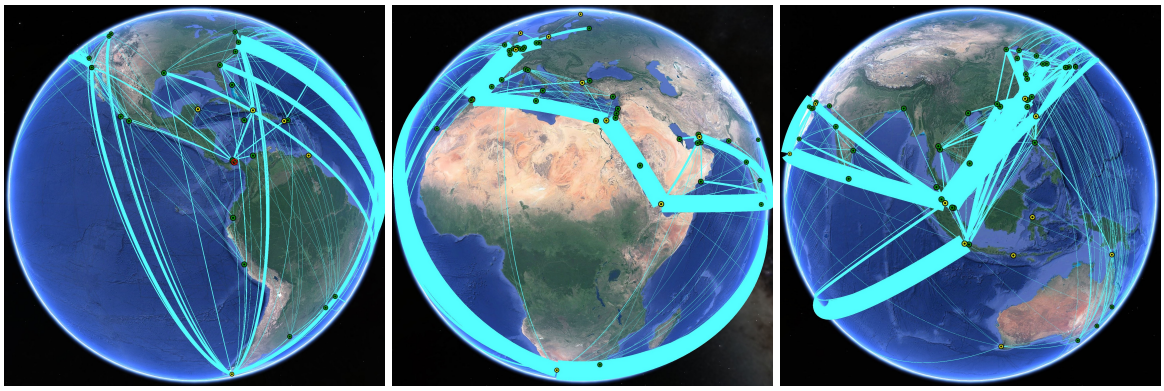
Scenario Results Table	
total transportation cost	$99,940,044,822
total shortfall/excess	0 TEUs
portion of road transportation	0.1%
portion of transportation between chokepoints	40%
ports with exhausted capacity	-

Table 4.6 provides the results of the interdiction. The increase in the total transportation cost is $6,355,097,051, which is about 6.8% of the base case cost. The reason for the increase is longer transportation routes from origin to destination. As seen in Figure 4.7, there is one road transportation arc (in green) that is utilized between Ashdod, Israel and Jeddah, Saudi

Arabia. But the amount of flow there is so small that none of the ports exhausts its capacity.



(a) Base case situation.



(b) Failure of the Panama Canal.

Figure 4.8. Comparison of the base case and the Panama Canal failure scenario.

The next scenario that we consider is an interdiction of the Panama Canal. Figure 4.8 shows the clear flow changes in this scenario. On the left, we recognize the decreased flow from Europe, but a significant increase from Cape of Good Hope to North America. Furthermore, the transition through the Strait of Magellan has become more important despite of the long way round for the most routes. In the middle, the huge increase of flow in the southern hemisphere between Asia and North America across the Indian Ocean and the Atlantic is clear to see. This is confirmed by the figures on the right, where the traffic through the Sunda Strait increased even more than in the previous scenario. A decrease of flow across the Pacific is also recognizable.

Table 4.7 shows the increased transportation cost in this scenario, which is here 10.3%

Table 4.7. Results of the interdiction of the Panama Canal.

Scenario Results Table	
total transportation cost	$103,200,517,821
total shortfall/excess	0 TEUs
portion of road transportation	0.8%
portion of transportation between chokepoints	40%
ports with exhausted capacity	Balboa, Panama Manzanillo International Terminal, Panama

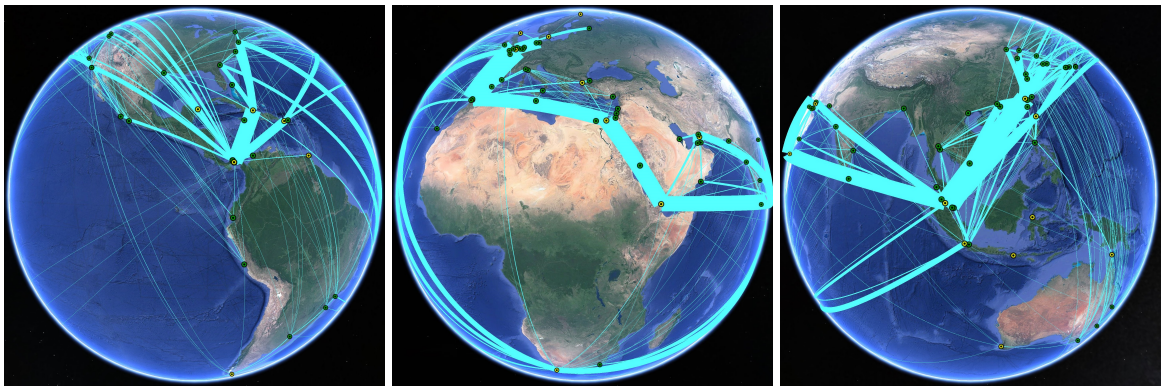
higher than in the base case. On both ends of the Panama Canal (Pacific and the Caribbean Sea) container ports are located in our model. These are the port of Balboa and the Manzanillo International Terminal. The increase of the road transportation portion of flow comes from increased cargo flow along the road connection between the two ports (until the port capacities are exhausted).

The last single interdiction scenario we consider is a blockade of the Strait of Malacca. The effects of the interdiction are hardly recognisable in Figure 4.9. In Asia (on the left), we observe a simple shift of traffic from the Strait of Malacca to the Sunda Strait. This result also appears in Table 4.8. The increase in the total transportation cost is approximately 2.2%, compared to the base case. There is no road transportation or ports with exhausted capacity result from the interdiction. The Sunda Strait seems to be a good substitute of the Strait of Malacca.

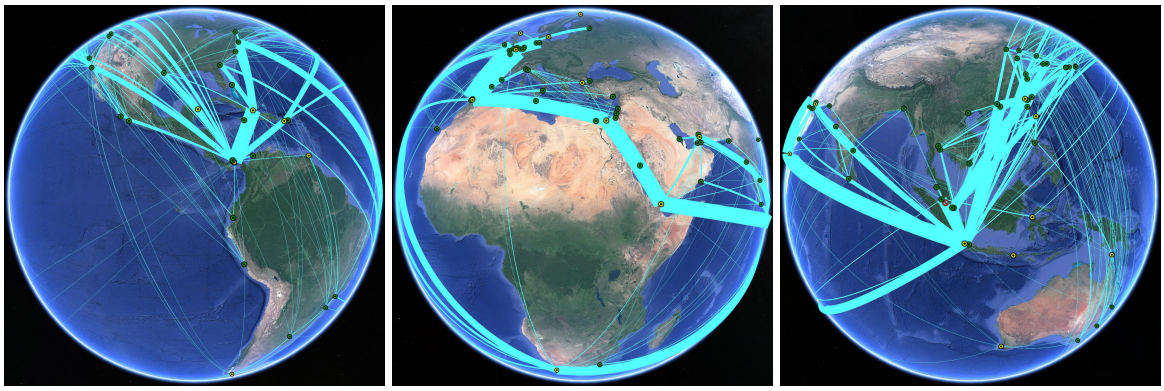
Table 4.8. Results of the interdiction of the Strait of Malacca.

Scenario Results Table	
total transportation cost	$95,682,578,400
total shortfall/excess	0 TEUs
portion of road transportation	0%
portion of transportation between chokepoints	40%
ports with exhausted capacity	-

We now consider a different question: which nodes, if interdicted, yield the biggest increase in system cost? To evaluate this, we exhaustively enumerate each possible scenario with exactly one interdicted port or maritime chokepoint. Table 4.9 lists the top ten nodes with greatest increase to the total transportation cost. The most expensive scenario is a failure of the port in Busan, which is the sixth biggest port in the world. If Busan is interdicted,



(a) Base case situation.



(b) Failure of the Strait of Malacca.

Figure 4.9. Comparison of the base case and the Strait of Malacca failure scenario.

all the cargo to and from this port is transported through the closest ports, which are far away and do not have sufficient capacity for this amount of redirected cargo. Therefore, three closest ports exhaust their capacity until all cargo demands are satisfied. This creates a lot of expensive road transportation and further increases the total cost of the scenario. Similar behavior is observed for scenarios associated with the interdiction of other ports in the table. The interdictions in the maritime chokepoints in the table create in general less road transportation, but cause instead long detours and thus increase the total cost.

There are some single-node interdiction scenarios that result in a shortfall of cargo for some nodes. This happens mostly to interdicted ports on islands that don't have adjacent ports on the same island. Then there is no alternative route for transportation. Since our model sets a very high penalty cost on having a shortfall, it is not realistic to add this artificial cost to

Table 4.9. Nodes with the highest transportation cost increase compared to the base case.

Rank	Closed node	Cost increase	Road portion	Exhausted ports
1	Busan	19.66%	5.79%	3
2	Strait of Gibraltar	15.45%	1.83%	3
3	Jawaharlal Nehru Port	11.50%	1.08%	3
4	Port of Shanghai	10.87%	4.14%	2
5	Panama Canal	10.27%	0.78%	2
6	Strait of Hormuz	9.16%	1.87%	0
7	New York	9.00%	2.18%	2
8	Bab-el-Mandeb	8.57%	0.11%	0
9	Shenzhen	8.16%	3.17%	2
10	Dubai	6.83%	1.61%	3

the total cost. But in reality, the penalty cost for undelivered cargo can still be very high. Table 4.10 provides an overview of the resulting shortfall in some scenarios.

Table 4.10. Single-node interdictions resulting in cargo shortfalls.

Rank	Closed node	Total scenario shortfall (in TEUs)
1	Tanjung Priok	4,460,060
2	Kaohsiung	2,865,533
3	Manila	2,148,180
4	Colombo	2,112,374
5	Marsaxlokk	1,352,232
6	Dublin	1,345,965
7	Tangier	1,237,929
8	Durban	1,124,715
9	Kingston	860,988
10	San Juan	764,339
11	Freeport	638,784
12	Las Palmas	555,737
13	Strait of Hormuz	4

We are now prepared to answer one of the initial questions of Chapter 4: Do the most "central" nodes from Chapter 3 also have the biggest impact on network flows in case of their failure? Table 4.11 compares the twenty nodes with the highest increase in total transportation cost to twenty most "central" nodes as defined by betweenness centrality. The table reveals that only eight nodes appear in both top twenty lists. Both sides of the

table show the high importance of the Strait of Gibraltar and the positions of the Panama Canal and Singapore are similar on both sides. Other than Busan, Strait of Hormuz, New York and Suez Canal, none of the other nodes appear in the top twenty lists, which are already $\frac{1}{6}$ of all available nodes. Previous observations showed, for example, that the Strait of Malacca, Dover Strait, Taiwan Strait and Mona Passage, all from the top ten "central" list, can be easily substituted by other nodes without much of additional cost. This demonstrates that looking at the node centrality is not enough to determine the importance of a node. Additional information about the network is required.

Table 4.11. Ranking of the most "expensive" and the most "central" nodes.

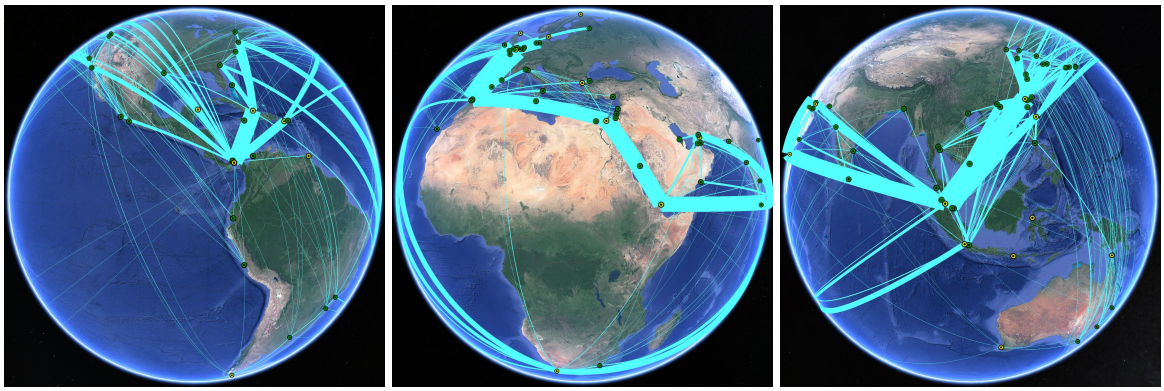
Rank	Highest cost increase nodes	Most "central" nodes
1	**Busan**	**Strait of Gibraltar**
2	**Strait of Gibraltar**	**Suez Canal**
3	Jawaharlal Nehru Port	**Bab-el-Mandeb**
4	Port of Shanghai	Strait of Malacca
5	**Panama Canal**	Bering Strait
6	**Strait of Hormuz**	**Panama Canal**
7	**New York**	Dover Strait
8	**Bab-el-Mandeb**	Taiwan Strait
9	Shenzhen	**Singapore**
10	Dubai	Mona Passage
11	**Suez Canal**	Cape of Good Hope
12	**Singapore**	Davis Strait
13	Savannah	Luzon Strait
14	Jeddah	Barents Sea
15	Qingdao	**Strait of Hormuz**
16	Houston	Windward Passage
17	Bandar Abbas	Sunda Strait
18	Tianjin	**Busan**
19	Cartagena	**New York**
20	Øresund	Dardanelles

4.3.2 Multiple Interdictions

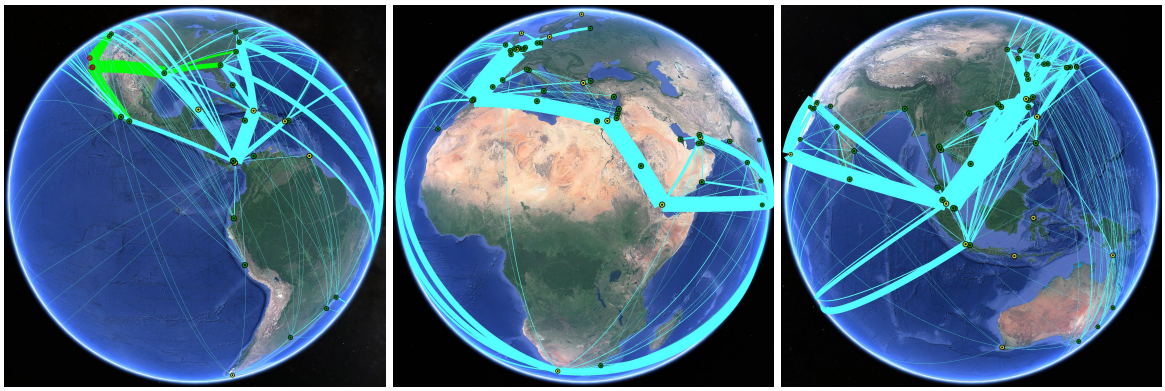
As an extension to the previous analysis, we take a closer look at scenarios with multiple port interdictions. We expect that, depending on the combination of nodes that are affected, the total impact of multiple interdictions might be greater than just the sum of the single scenario impacts. If two closed ports are in the same area, they serve as substitutes for

one another. This way, farther ports need to process cargo for multiple ports and the transportation cost might explode. In this subsection we only consider port failures, that is we assume that maritime chokepoints are always available.

First, we evaluate a specific scenario that reminds of the West Coast port labor lockout in 2002. For the scenario, we close three major U.S. ports: Long Beach, Los Angeles and Oakland. These ports constitute a major portion in North America's West Coast logistics network. Since the ports have a big total cumulative throughput and they are located close to each other, we expect a lot of road transportation and therefore high cost.



(a) Base case situation.



(b) Failure of the ports in Long Beach, Los Angeles and Oakland.

Figure 4.10. Comparison of the base case and the West Coast ports failure scenario.

The resulting flow in Figure 4.10 confirms our expectations. On the left, it shows many thick road transportation arcs across North America. All West Coast ports in Canada and Mexico are busy supporting the three ports. Even the ports at the east coast are involved in

processing the cargo of the interdicted ports. Both other perspectives indicate an increase in sea transportation routes between North American's East Coast ports and both continents Europe and Asia.

Table 4.12. Results of the interdiction of the West Coast ports.

Scenario Results Table	
total transportation cost	$169,079,946,490
total shortfall/excess	0 TEUs
portion of road transportation	8.3%
portion of transportation between chokepoints	36%
ports with exhausted capacity	Houston, Lazaro Cardenas, Manzanillo, Metro Vancouver, Savannah, Seattle

The results in Table 4.12 emphasize the visual observation from the map. The road transportation flow has increased to 8.3%. Due to this fact, the total transportation cost in the scenario is 80.7% higher than in the base case. This increase in value is caused by just about 3% of the ports. Six container ports in Canada, Mexico and the U.S. exhaust their capacities while supporting the interdicted ports. Additional port failures in North America would most likely cause shortfalls because the cumulative capacity would be exhausted there.

Finding meaningful combinations of failing ports is a challenging task, because too many possible combinations exist. Therefore, we restrict our analysis to exactly two concurrent failing ports and evaluate every possible combination to produce a ranking similar to that for single interdictions. Having 94 container ports in the model, this means considering a total of 4,371 scenarios which takes about six days of total computation time.

Table 4.13 lists the twenty two-port combinations that cause the highest increase to transportation costs. The worst scenario by far is the scenario involving the failure of Long Beach and Los Angeles ports, which results in a 52.2% cost increase. The second and third positions, involving scenarios of failures in the Malay Peninsula and in China, are considerably less costly. In the further positions, the cost decrease smoothly towards the end of the table.

Although the highest value is caused by U.S. ports, the port of Busan is represented

Table 4.13. Highest cost increase for double interdictions.

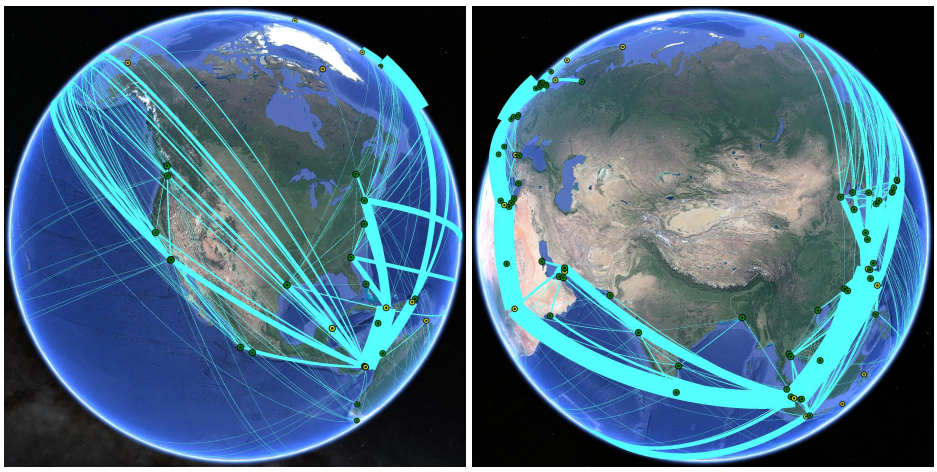
rank	closed node	cost increase	road portion	exhausted ports
1	Long Beach, Los Angeles	52.20%	5.98%	6
2	Port Klang, Singapore	36.38%	5.71%	6
3	Guangzhou, Shenzhen	33.25%	4.17%	2
4	Ningbo, Port of Shanghai	32.95%	5.35%	4
5	New York, Savannah	30.44%	3.46%	4
6	Busan, Port of Shanghai	29.90%	8.75%	5
7	Singapore, Tanjung Pelepas	29.31%	5.02%	5
8	Busan, New York	28.67%	6.71%	5
9	Busan, Shenzhen	27.82%	7.61%	5
10	Busan, Dalian	27.69%	6.02%	3
11	Busan, Kwangyang	27.53%	6.34%	3
12	Busan, Tianjin	26.73%	7.13%	4
13	Busan, Dubai	26.50%	6.45%	6
14	Busan, Incheon	25.71%	5.94%	3
15	Busan, Singapore	25.21%	8.61%	5
16	Busan, Savannah	25.11%	5.96%	4
17	Busan, Qingdao	24.89%	6.42%	4
18	Busan, Houston	24.70%	5.68%	3
19	Hampton Roads, New York	23.78%	2.89%	3
20	Busan, Cartagena	23.51%	5.28%	4

disproportionately in the first twenty positions. The reason is that this port caused the highest cost increase in the single interdiction scenarios. It is surprising that the worst two-port interdictions are not only ports that are located near to each other, but also include some combinations of ports from different continents. The highest road portion value of all scenarios is 8.75% and at most six ports exhaust their capacities in each scenario.

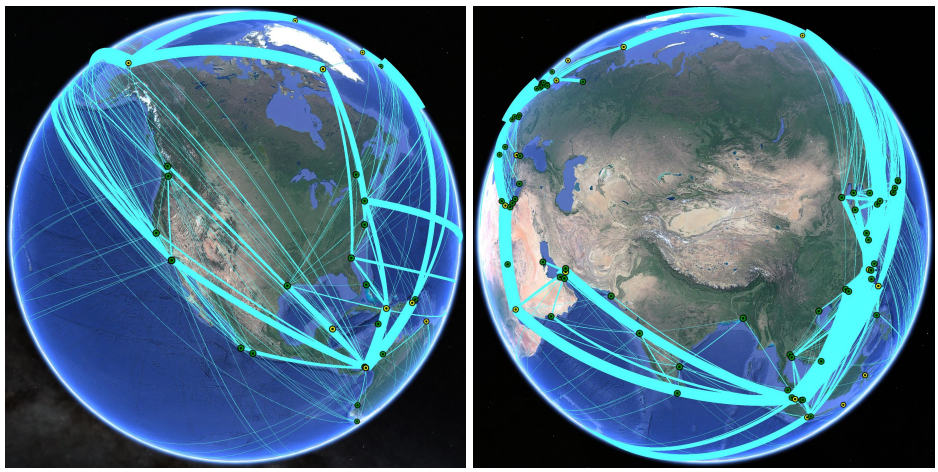
4.3.3 Arctic Sea Routes

The last analysis in this chapter concerns the arctic sea routes. As mentioned earlier, the Northwest Passage and the Northeast Passage are usable only for a few months of a year because of the ice. Therefore, only a few ships per year choose these routes and therefore we have set an artificial capacity there at 18 ships per year. But due to climate change and melting arctic ice, the availability of these routes could increase. Then they would become more attractive for shipping companies, because in some cases they are shorter

than the routes used today. To investigate this statement, we increase the capacity of the arctic routes and observe the resulting impact to the global transportation system. As an example, we decide for the current throughput of the port of Los Angeles and implement it at the common entrance of both arctic routes—the Bering Strait.



(a) Base case situation.



(b) Increased capacity on the arctic sea routes.

Figure 4.11. Comparison of the base case and the increased arctic traffic scenario.

The resulting cargo traffic is presented in Figure 4.11. Not surprisingly, we observe that both arctic arcs are utilized much more than in the base case. The figures in the left columns show, that the flow from Asia to the Panama Canal has decreased. In general, there exists less cargo flow between the Panama Canal and the U.S. East Coast. The East Coast is now

supplied better by the Northwest Passage than before. The perspective on the right column substantiates the reduced flow between Asia and the Mediterranean Sea. A great portion of the flow from Asia is now directed to the Bering Strait and from there further to the arctic routes.

Table 4.14. Results of the increased capacity in the arctic sea routes.

Scenario Results Table	
total transportation cost	$90,178,602,508
total shortfall/excess	0 TEUs
portion of road transportation	0%
portion of transportation between chokepoints	35%
ports with exhausted capacity	-

Table 4.14 shows the resulting numbers of the last scenario. The total transportation cost is about 3.6% lower than in the base case scenario. This is a reasonable saving of money and fuel for global transportation, if we assume the same transportation cost through the arctic routes as for all other routes. The result also shows that the scenario has exhausted the capacity of the Bering Strait. This means that the arctic routes most likely would carry even more cargo in the optimal solution, if we allow a higher capacity in the Bering Strait.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 5:
Summary

5.1 Conclusions

This thesis considers two perspectives on the global maritime transportation network—one informed by network science and another informed by a multi-commodity network flow problem.

The network science perspective allows us to study general characteristics based on graph connectivity. Applying centrality measures, we identify container ports and maritime chokepoints that are reasonable candidates for the most important nodes of the network. We observe that the betweenness centrality fits our analysis well and if some of the predicted most "central" nodes fail, we indeed see the highest cost increases. The ranking positions in terms of betweenness centrality also conformed closely to the ones of the highest cost increases in case of failure.

However, the majority of the most "central" nodes were not the most important to the network for our application. The betweenness centrality only identified important nodes based on the topology of the network. However, our network contains additional information, such as flow and node capacities, which change the importance of nodes in case of failures. For port failures in our scenarios, the existence of nearby ports with sufficient capacities is essential for mitigating excessive costs. For closed maritime chokepoints, it is important to have an alternative route that is not a long detour. If these requirements are not met, the cost can grow very fast. The centrality definitions are not able to take such details into account.

The Multi-Commodity Linear Optimization Model and its code implementation allow quick creation, computation and visualization of arbitrary scenarios with any combination of node interdictions. This way, plausible potential real world scenarios can be produced and analyzed very fast. Due to the generic nature of the linear model, it can be utilized with any model of the global maritime transportation network. Additional constraints, like road route capacities can be easily implemented.

Studying the single node interdiction scenarios, we identified the port of Busan, South Korea as the node with by far the biggest transportation cost increase in case of its failure. The cost increase is mostly based on the additional road transportation. The second biggest increase is caused by the Strait of Gibraltar, because its blockade would cause large detours for the cargo between Europe and China. Other less expensive scenarios have different reasons for the cost increase and need to be analyzed separately for more precise understanding of the results.

The results of the double interdiction scenarios show as the worst scenario the simultaneous failure of the ports in Los Angeles and Long Beach due to their current importance for intercontinental trade. Most of the other results involve ports located in Asia. But the ranking also shows that the interdictions don't need to happen on the same continent to cause a big increase in transportation cost.

Finally, the results of the small arctic routes study, which doesn't contain any interdictions, show that these routes have a potential to dramatically decrease global transportation cost, once it becomes possible to use these routes in the same manner as today's most common routes.

5.2 Future Work

Our model of the global maritime transportation network is limited to 94 major global container ports from 58 different countries. But additional smaller ports exist worldwide, and their cumulative throughput constitutes an important portion of the global container traffic. Many countries with smaller ports are not represented in our model. Including additional ports in the model would make it more complex but also more realistic.

Another possible extension could be to implement additional commodities, like different types of bulk cargo. This would increase the computation time, but on the other hand would provide better results. Another idea is to differentiate between different types of ports, since not every port is able to handle all sizes of vessels.

For the purpose of cargo transportation, the model consists of sea and road network, since the data on route distances is relatively easy to collect. But in many countries in the real world cargo is also transported by railway. To gain more precise results on detour cost and

allow additional routes in the network, railway network data should be collected. Having the data only for one continent would already improve the results for node failure scenarios in that region.

Another potential area for future work could be to add time constraints for cargo delivery. Some types of cargo require just in time delivery or have specific requirements for maximum transport duration. Therefore, in some cases it would not be feasible to choose any of the available detours. If the cargo is not delivered by the required time, the demand will not be met. In this case, it would be reasonable to analyze different scenarios to identify the portions of cargo, that cannot be delivered by time.

THIS PAGE INTENTIONALLY LEFT BLANK

List of References

[1] T. S. U. of New York. (2016). Trade and globalization. [Online]. Available: http://www.globalization101.org/trade-introduction. Accessed February 12. 2017.

[2] T. Huber, ""Global Ports" in der maritimen Transportwirtschaft. Akteursbasierende Bewertung des weltweiten Netzwerks von Hafenstandorten ["global ports" in the maritime transport sector, a player based assessment of the global network of port locations]," Ph.D. dissertation, Fakultät für Philosphie, Kunst-, Geschichts- und Gesellschaftswissenschaften der Universität Regensburg, Regensburg, Germany, 2014.

[3] S. Bruns, K. Petretto, and D. Petrovic, *Maritime Sicherheit [maritime security]*. Springer-Verlag, 2013.

[4] A. Khouri. (2015). West Coast ports facing shutdown in labor dispute. *Los Angeles Times*. [Online]. Available: http://www.latimes.com/business/la-fi-port-shutdown-20150207-story.html. Accessed May 29. 2017.

[5] S. S. Cohen, "Economic impact of a West Coast dock shutdown," *University of California at Berkeley*, p. 1, 2002.

[6] R. Guimera, S. Mossa, A. Turtschi, and L. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *Proceedings of the National Academy of Sciences*, vol. 102, no. 22, pp. 7794–7799, 2005.

[7] P. Kaluza, A. Kölzsch, M. T. Gastner, and B. Blasius, "The complex network of global cargo ship movements," *Journal of the Royal Society Interface*, vol. 7, no. 48, pp. 1093–1103, 2010.

[8] C. Wang and J. Wang, "Spatial pattern of the global shipping network and its hub-and-spoke system," *Research in Transportation Economics*, vol. 32, no. 1, pp. 54–63, 2011.

[9] T. G. Martagan, B. Eksioglu, S. D. Eksioglu, and A. G. Greenwood, "A simulation model of port operations during crisis conditions," in *IEEE Simulation Conference (WSC), Proceedings of the 2009 Winter*, 2009, pp. 2832–2843.

[10] C. Madhusudan and G. Ganapathy, "Disaster resilience of transportation infrastructure and ports – an overview," *International Journal of Geomatics and Geosciences*, vol. 2, no. 2, p. 443, 2011.

[11] E. Miller-Hooks, X. Zhang, and R. Faturechi, "Measuring and maximizing resilience of freight transportation networks," *Computers & Operations Research*, vol. 39, no. 7, pp. 1633–1643, 2012.

[12] E. D. Pidgeon, "Modeling the effects of a transportation security incident upon the marine transportation system," Master's thesis, Monterey, California. Naval Postgraduate School, 2008.

[13] L. A. Bencomo, "Modeling the effects of a transportation security incident on the commercial transportation system," Master's thesis, Monterey, California. Naval Postgraduate School, 2009.

[14] G. Brown, W. Carlyle, J. Salmeron, and K. Wood, "Analyzing the vulnerability of critical infrastructure to attack, and planning defenses," 2005.

[15] K. Wood, G. Brown, M. Carlyle, and J. Salmerón, "Defending critical infrastructure," *Interfaces*, vol. 36, pp. 530–544, 2006.

[16] J. P. Babick, "Tri-level optimization of critical infrastructure resilience," Master's thesis, Monterey, California. Naval Postgraduate School, 2009.

[17] J. J. Onuska, "Defending the Pittsburgh waterways against catastrophic disruption," Master's thesis, Monterey, California. Naval Postgraduate School, 2012.

[18] O. R. Garcia Olalla, "Assessing the resilience of global sea routes," M.S. thesis, Monterey, California. Naval Postgraduate School, 2012.

[19] C. F. de la Cruz, "Defending the maritime transport of cargo for the Hawaiian Islands," Master's thesis, Monterey, California. Naval Postgraduate School, 2011.

[20] W. Wenke, "Assessing the operational resilience of the Port of Anchorage: Recommendations for investment and implications for policy," Master's thesis, Monterey, California. Naval Postgraduate School, 2015.

[21] J.-P. Rodrigue, "Straits, passages and chokepoints: a maritime geostrategy of petroleum distribution," *Cahiers de géographie du Quebec*, vol. 48, no. 135, pp. 357–374, 2004.

[22] B. I. Singapore, "The 7 biggest oil chokepoints in the world — seen like never before," November 2016, accessed February 19. 2017. Available: http://www.businessinsider.sg/biggest-oil-chokepoints-2016-11/9

[23] CIA, "The world factbook," February 2017, accessed February 2017. Available: https://www.cia.gov/library/publications/the-world-factbook/

[24] J.-P. Rodrigue, "The geography of transport systems," August 2013, accessed February 21. 2017. Available: http://people.hofstra.edu/geotrans/eng/ch3en/conc3en/img/Map_main_shipping_routes.png

[25] W. Komiss and L. Huntzinger, "The economic implications of disruptions to maritime oil chokepoints," *Center for Naval Analysis*, 2011.

[26] J. H. Noer and D. Gregory, "Chokepoints: Maritime economic concerns in Southeast Asia." DTIC Document, Tech. Rep., 1996.

[27] R. Mitchell, *Web Scraping with Python: Collecting Data from the Modern Web*. O'Reilly Media, Inc., 2015.

[28] J. Elkner, "Getting down with CSS," January 2017, accessed February 26. 2017. Available: http://www.openbookproject.net/tutorials/getdown/css/lesson4.html

[29] R. Gera, "MA4404: structure and analysis of complex networks," March 2017, accessed March 29. 2017. Available: http://faculty.nps.edu/rgera/MA4404.html

[30] M. Newman, *Networks: An Introduction*. OUP Oxford, 2010.

[31] M. Bastian *et al.* Gephi. [Online]. Available: https://gephi.org. Accessed January 30. 2017.

[32] J. D. Hunter. Matplotlib. [Online]. Available: http://matplotlib.org. Accessed February 10. 2017.

[33] P. Basu, R. Sundaram, and M. Dippel, "Multiplex networks: A generative model and algorithmic complexity," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. ACM, 2015, pp. 456–463.

[34] L. G. Jeub, P. Balachandran, M. A. Porter, P. J. Mucha, and M. W. Mahoney, "Think locally, act locally: Detection of small, medium-sized, and large communities in large networks," *Physical Review E*, vol. 91, no. 1, p. 012821, 2015.

[35] A. Clauset, "Network analysis and modeling," university lecture, University of Colorado Boulder, 2013.

[36] D. L. Alderson, G. G. Brown, and W. M. Carlyle, "Assessing and improving operational resilience of critical infrastructures and other systems," in *Bridging Data and Decisions*. INFORMS, 2014, pp. 180–215.

[37] "Container port capacity study," The Tioga Group, Inc., December 2010.

[38] A. Council *et al.*, "Arctic marine shipping assessment 2009," 2009. Available: http://library.arcticportal.org/id/eprint/1400

[39] "Situation and analysis of Panama Channel's capacity," Panama Canal Authority, 2012.

[40] W. Murray, "Economies of scale in container ship costs," United States Merchant Marine Academy, 2015. Available: http://www.jstor.org/stable/20053805

[41] R. Asariotis, H. Benamara, J. Hoffmann, M. Misovicova, E. Nunez, A. Premti, B. Sitorus, V. Valentine, and B. Viohl, "Review of maritime transport, 2010," Tech. Rep., 2010.

[42] P. Zeihan, *The accidental superpower: the next generation of American preeminence and the coming global disorder*. Twelve, 2014.

[43] W. E. Hart, J.-P. Watson, D. L. Woodruff, J. Siirola, and C. Laird. Pyomo. [Online]. Available: http://www.pyomo.org. Accessed April 30. 2017.

[44] W. E. Hart, J.-P. Watson, and D. L. Woodruff, "Pyomo: Modeling and solving mathematical programs in Python," *Mathematical Programming Computation*, vol. 3, no. 3, pp. 219–260, 2011.

[45] W. E. Hart, C. Laird, J.-P. Watson, and D. L. Woodruff, *Pyomo–optimization modeling in python*. Springer Science & Business Media, 2012, vol. 67.

[46] IBM. ILOG CPLEX Optimizer. [Online]. Available: https://www-01.ibm.com/software/commerce/optimization/cplex-optimizer. Accessed April 30. 2017.

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California